Decision Sciences Institute

Scholars • Educators • Problem Solvers

Northeast Decision Sciences Institute

52nd Annual Conference

March 30 – April 1, 2023

Washington, DC Metro Area

RISK, RESILIENCE AND RECOVERY IN TURBULENT BUSINESS ENVIRONMENTS

NEDSI
NORTHEAST DECISION SCIENCES INSTITUTE

UNIVERSITY OF THE DISTRICT OF COLUMBIA 1851

NEDSI

NORTHEAST DECISION SCIENCES INSTITUTE

NEDSI 2023

Northeast Decision Sciences Institute

52nd Annual Conference

RISK, RESILIENCE AND RECOVERY IN TURBULENT BUSINESS ENVIRONMENTS

HOST INSTITUTION

School of Business and Public Administration

University of the District of Columbia

UNIVERSITY OF THE DISTRICT OF COLUMBIA 1851

March 30 – April 1, 2023

Washington, DC Metro Area

UNIVERSITY OF THE DISTRICT OF COLUMBIA
SCHOOL OF BUSINESS AND PUBLIC ADMINISTRATION

Developing the Business Leaders of Tomorrow

The School of Business and Public Administration prepares students to become adaptive, innovative and ethical leaders with a global perspective and real-world solutions to issues that challenge public and private organizations.

Degree Programs

Bachelor of Business Administration (BBA) in Accounting*
Bachelor of Business Administration (BBA) in Finance*
Bachelor of Business Administration (BBA) in Business Management **
Master of Business Administration (MBA)***

** with concentration in Logistics and International Trade*
*** with concentrations in Logistics and International Trade, Management Information Systems, and Marketing*
**** with concentrations in Public Management and International Business*

More Information

4200 Connecticut Avenue, NW, Washington, DC 20008
(202) 274-7000 | sbpa@udc.edu | **udc.edu/sbpa**

AACSB ACCREDITED

TABLE OF CONTENTS

WELCOME

Dear NEDSI Members and Guests,

As the Chief Academic Officer of the University of the District of Columbia (UDC), the host institution of the 52nd Annual Conference of the Northeast Decision Sciences Institute (NEDSI), I am pleased to welcome you to this special occasion. As the nation's only public, urban land-grant, emerging research HBCU, UDC is committed to fostering academic excellence, collaborative research activities, and community engagement. Our recent announcement of becoming an AACSB-accredited university is evidence of our pursuit of excellence. These are significant reasons why UDC is excited to participate in this vital conference that brings together a diverse group of scholars, practitioners, and students in various business and education fields.

This year's conference theme, "Risk, Resilience, and Recovery in Turbulent Business Environments," illuminates the importance of decision sciences and data-driven decision-making across disciplines to achieve the best outcomes—especially needed in today's dynamic business landscape. During the three days of the Conference, you can engage in stimulating presentations and discussions, share innovative ideas, and collaborate on cutting-edge research and pedagogical techniques.

The University recognizes the vital role that NEDSI plays in advancing knowledge and research and is honored to host this prestigious Conference. I wish you a productive and successful event on behalf of UDC's administration, faculty, staff, students, trustees, alums, and the larger community.

Sincerely,

Lawrence T. Potter, Jr, Ph.D.
Chief Academic Officer
University of the District of Columbia

WELCOME

On behalf of the School of Business and Public Administration (SBPA), at the University of the District of Columbia, I am delighted to welcome you to our town for the 2023 Northeast Decision Sciences Institute (NEDSI) Annual Meeting, in Washington DC.

While watching from the sidelines, I have been observing, and I truly appreciate the hard and the delicate work of the organizing committee to piece together the intricate parts of the program, to craft a remarkable conference program.

We are honored and excited to be the host institution for the prestigious NEDSI conference, this year. It adds to and complements our AACSB accreditation, earned just a few weeks earlier. All these are consistent with our commitment to SBPA's core values of Leadership Skills, Professionalism, Business Expertise, and Strategic Mindset. The SBPA is now the only AACSB accredited public institution of higher learning in the Nation's Capital.

During the past few years, the Business School has transformed into a top-quality School, with innovative programs, and receiving the designation of the best value quality business program. In line with our Mission and Vision, and meeting the AACSB expectations on the social impact, we take pride in offering innovative and cutting-edge programs to meet and to exceed those expectations. The Social Motivation Approach for Rehabilitation Through Educational Robotics (SMARTER), The Paving Access Trails to Higher Security (PATHS), and the VITA Tax Clinic are among few creative programs that have meaningful impact on the advancement and upward mobility of the community.

We hope you will enjoy the 2023 conference and your stay in the Nation's Capital. Washington, DC, is a cosmopolitan mecca and one of the most exciting and prominent cities in the world. The city embodies vibrant multicultural communities and environment, renowned national museums, world leading art and cultural centers, and celebrated ethnic and cultural magnets. Washington provides exceptional prospects and excellent opportunities for research, consulting, and grants at local, district and federal levels.

Mohamad Sepehri
Dean – School of Business and Public Administration
University of the District of Columbia

WELCOME

Welcome to NEDSI 2023!

52nd Annual Conference of Northeast Decision Sciences Institute

On behalf of the entire NEDSI Board and Organizing Committee, we would like to welcome you to the 52nd Annual Conference of NEDSI. The theme of the conference is "Risk, Resilience and Recovery in Turbulent Business Environments." Through the efforts of numerous scholars, thought leaders, and generous volunteers, we are excited to offer you once again a variety of paper sessions, workshops, panels, and consortia that we hope will engage you in professional development and networking opportunities.

The board and organizing committee have worked very hard over the past few months to pull together a great 2023 conference. In addition to our sincere gratitude to the entire NEDSI board, we are very grateful to Dr. Lawrence T. Potter, Jr., Chief Academic Officer and Dr. Mo Sepehri, Dean – School of Business and Public Administration (SBPA) at the University of the District of Columbia (UDC), without whom it would not have been possible to host the conference in Washington DC. SBPA-UDC is the host institute for NEDSI 2023.

The venue for NEDSI 2023 is the Washington DC metro area. The conference hotel is Crystal Gateway Marriott located in Arlington, VA, just across the district line and close to all the Washington DC attractions including Smithsonian Museums, White House, Washington Monument, Capitol Hill, National Mall, and many more! All attractions are easily accessible via Crystal City Metro station which is connected from the hotel lobby.

On behalf of the NEDSI board, I wish you all the best and hope you find NEDSI 2023 fruitful and enjoyable.

Best regards,

Amit Arora, Ph.D.
Program Chair NEDSI 2023
Associate Professor of Supply Chain Management
University of the District of Columbia

KEYNOTE SPEAKER

Rhom Erskine is Vice President, Global Diversity & Inclusion at Lockheed Martin, with specific responsibility for diversity and inclusion vision, strategic direction, employee engagement, and equal opportunity compliance for the company and its 110,000+ employees. In this role, he works closely with business area and HR executives, talent acquisition/management leaders, communications, and external partners to drive diversity and inclusion thought leadership and impact.

Rhom is a senior consultative leader with over 20 years of experience in large, mid-size, and small business environments, where he has demonstrated success in strategic Diversity & Inclusion (D&I) planning and growth, D&I Council leadership, Business Resource Groups, change management, employee engagement strategies, diversity recruiting, and product/program management. Having spent the early part of his career in marketing, software product development, and sales strategy, he also brings strong line-of-business experience to complement award-winning D&I experience.

Prior to this role, Rhom was Director, Global Diversity & Inclusion at Lockheed Martin's Aeronautics business area. Before joining Lockheed Martin, he served as Director of Diversity & Inclusion at McKesson Corporation, a Fortune 10 pharmaceutical distribution and medical technology company. Previous roles also include Director of Diversity, Inclusion, and Multicultural at Affairs at Harland Clarke Holdings Corporation, where his focus extended internationally as well, leading the company's Global Advisory Team, which functioned as an internal knowledge repository, providing guidance to project teams and individuals to reduce the cultural learning curve during international acquisitions.

Rhom supports a number of community and professional organizations, including past/present Board of Directors positions with Advancing Minorities Interest in Engineering (AMIE), the National Black MBA Association's Atlanta Chapter, the Clark Atlanta University School of Business, the Center for the Visually Impaired , and Zoo Atlanta.

A native of Huntsville, Alabama, Rhom holds a B.A. in Psychology from Morehouse College, an MBA with a Marketing concentration from Clark Atlanta University, and has received executive training in Diversity & Inclusion from Georgetown University and the American Institute for Managing Diversity. He has been selected as one of the Top 40 Under 40: Georgia’s Best and Brightest by Georgia Trend Magazine, as well as a member of the Top 100 Under 50 Emerging and Executive Leaders by Diversity MBA Magazine. He and his wife Danielle reside in Atlanta, GA and are the proud parents of three sons.

"FIRESIDE CHAT" SPEAKERS

Michael Kenney, VP-Operational Risk, Freddie Mac
As Vice President of Operational Risk, Michael leads the first line of defense risk management for financial crimes, privacy and information security, business resiliency, vendor management and compliance with regulatory requirements. Mike and his team establish the risk direction by aligning the Multifamily Governance Framework with the business operating model. He continually improves governance by understanding the multifaceted drivers that affect risk environment.

Thadi Murali, Managing Principal, Capco Consulting
Thadi Murali is a Managing Principal with Capco, a management consultant with 18+ years of experience. He provides consulting services to the world's top financial institutions, with a focus on Risk, Compliance, and Data Management across Banking and Asset Management. He has helped clients in the financial industry on both large and small transformation engagements, design and implement solutions to address risk (Enterprise and Operational), regulatory challenges in banking (Basel II, III, Dodd-Frank), and data management (Regulatory Reporting, CCPA and GDPR). He is a thought leader. Besides being a sought after consultant, he is an industry thought leader, who has published many articles and spoken in many industry forums. His article on Data Resiliency shows how taking a data-centered approach strengthens an organization's ability to plan, anticipate, detect, correct, and build a sustainable operational resilience culture.

Kelly Siu, Privacy Risk Program Manager, Meta (fka Facebook)
Kelly Siu holds a Bachelor of Science in Computer Information Systems and Master of Science in Management from American University. Kelly is a member of the Association of Certified Fraud Examiners to reduce the incidence of fraud and white-collar crime.
Upon graduating from American University, Kelly relocated to the San Francisco Bay Area to join IBM-Silicon Valley Laboratory as a software engineer for four years. Then she joined KPMG LLP's Technology Risk & Assurance Practice in Washington DC. Kelly was also at Freddie Mac and wore many heads while leading a risk management team. She is a subject matter expert in third party data sharing risk, privacy, and controls testing. Kelly is currently at Meta (fka Facebook) as a Privacy Risk Program Manager. Fun Facts: Kelly was born in Hong Kong and raised in Washington DC. She is the mother of 2 senior rescue dogs and volunteers at Operation Paws for Homes. She enjoys eating her way across the globe and traveling on a shoestring budget.

NJIT

Martin Tuchman School of Management
New Jersey Institute of Technology

Transforming Business Education

Online Accelerated Bachelor's Degree

Challenging the status quo of traditional institutions, NJIT is launching a new and innovative program that is flexible, accelerated and completely online. It is designed to meet the needs of adult learners and working professionals.

Graduate Certificate in Business Analytics

Responding to the rapid growth in demand for analytics professionals, this program is designed to equip business professionals with the quantitative and technical skills required for robust careers in the data-driven world of business.

Ph.D. Program in Business Data Science

Educating the next generation of data scientists, the program is one of the first of its kind in the country for business and management. The curriculum integrates business analytics and management systems theory with statistics, computing science and engineering.

Entrepreneurship Accolades

Placed among The Princeton Review's Top 50 Undergraduate Schools for Entrepreneurship Studies!

NJIT is ranked No. 41 among the Top 50 Undergraduate Schools for Entrepreneurship Studies for 2020. This is a recognition of NJIT's strengths in entrepreneurship curriculum, scholarly research, funding and mentorship for students, university-affiliated startups and service in this discipline.

Statewide Recognition for Our Faculty in Entrepreneurship Research and Teaching

We offer 17 credit-bearing entrepreneurship courses, currently to more than 450 students, as well as both major and minor entrepreneurship degrees. Our faculty were awarded the Bright Idea award for their outstanding publication, "The Role of Innovation Ecosystems and Social Capital in Startup Survival."

We offer internationally accredited and high value academic undergraduate and graduate programs that are business focused, technology integrated and experientially driven in various delivery modes.

For More Information: Email: **management@njit.edu** Visit: **management.njit.edu**

AACSB ACCREDITED

A Top 100 National University
-U.S. News & World Report

UNIVERSITY HEIGHTS • NEWARK, NJ 07102 • NJIT.EDU

The Decision Sciences Institute

The Decision Sciences Institute (DSI) is a professional organization of academicians andpractitioners interested in the application of quantitative and behavioral methods to the problems of society.
Through national, international and regional conferences, competitions and publications, the DSI provides an international forum for presenting and sharing researchin the study of decision processes across disciplines. The DSI also plays a vital role in the academic community by offering professional development activities and job placement services.
Five regional subdivisions in the United States, as well as regions representing Europe, Asia-Pacific and the Indian subcontinent, operate within the DSI. Each region has its ownelected officers and representative on the Board of Directors and holds annual meetings.
The DSI, an independent nonprofit educational organization, is located in Houston, Texas, where it receives extensive support from the C.T. Bauer College of Business at the University of Houston.

Northeast Decision Sciences Institute

The Northeast Decision Sciences Institute (NEDSI) is one of five regions in the AmericasDivision of the professional society, the Decision Sciences Institute. This region encompasses the Northeastern United States. NEDSI holds an annual regional meeting each spring that features presentations of original research papers, Ph.D. and new faculty development seminars, a placement service, case writing and other interesting innovationsin the fields of accounting, business education, finance, HRM, management, marketing, MIS/DSS/expert systems and IT, organizational behavior, operations management, strategic management and supply chain management, among others. Awards for “Best Paper” in several categories are given each year.

NEDSI 2023 CONFERENCE SPONSORS

University of the District of Columbia
School of Business and Public Administration
Host Institution
Grand Platinum Sponsor

New Jersey Institute of Technology
Martin Tuchman School of Management
Platinum Sponsor

Boston College
Carroll School of Management
Gold Sponsor

Georgia Institute of Technology CIBER
Gold Sponsor

Indiana University CIBER
Gold sponsor

Penn State Harrisburg
School of Business Administration
Gold Sponsor

University of Maryland Center for Global Business
Gold Sponsor

University of Rhode Island
College of Business
Gold Sponsor

Penn State Great Valley
School of Business Administration
Bronze Sponsor

The Case Center
Advertisement

NEDSI 2023 CONFERENCE EXHIBITORS

Palgrave Macmillan

Georgia Tech Scheller College of Business

Center for International Business Education and Research

The Georgia Tech Center for International Business Education and Research

Helping to ensure the long-term international economic competitiveness of the United States through support of research, business education initiatives, and community outreach activities.

ciber.gatech.edu
Georgia Institute of Technology
800 W. Peachtree St., N.W.,
Atlanta, GA, 30308-1149 USA
ciber@scheller.gatech.edu

Education | Research | Outreach

NEDSI BOARD OF DIRECTORS 2022-2023

President	**Dinesh R. Pai**, Penn State Harrisburg
Immediate Past President	**Gang Li**; **Effie Stavrulaki,** Bentley University
President-Elect	**Theologos Homer Bonitsis**, New Jersey Institute of Technology
Program Chair 2023	**Amit Arora**, University of the District of Columbia
Program Co-Chair-Elect 2024	**Neil Desnoyers**, St. Joseph's University
Program Co-Chair-Elect 2024	**Xiangrong Liu**, Bridgewater State University
V.P. Member Engagement	**Afrooz Moatari-Kazerouni**, Widener University
V.P. Promotional Activities	**Mehmet G. Yalcin**, University of Rhode Island
V.P. Technology	**EunSu Lee**, New Jersey City University
Treasurer	**Xiangrong Liu**, Bridgewater State University
Secretary	**Andres Fortino**, New York University
Archivist	**Jennifer A. Swanson,** Stonehill College

OFFICERS AT LARGE 2022 - 2023

Oya I Tukel, New Jersey Institute of Technology
Neil Desnoyers, St. Joseph's University
Linda Boardman Liu, Boston College
Douglas Hales, University of Rhode Island

PRESIDENTIAL ADVISORY COUNCIL 2022-2023

Eric W. Stein, Penn State Great Valley
Douglas Hales, University of Rhode Island
Neset Hikmet, University of South Carolina
Pedro M. Reyes, Baylor University
Joy Field, Boston College

ADVISORS AT LARGE 2022-2023

Effie Stavrulaki, Bentley University
Gang Li, Bentley University
Fariborz Partovi, Drexel University

PROGRAM TRACK CHAIRS

Accounting, Finance and Economics
Theologos Homer Bonitsis, New Jersey Institute of Technology | theologos.bonitsis@njit.edu

Big Data, Analytics and Knowledge Management
EunSu Lee, New Jersey City University | elee3@njcu.edu

Cyber Security, IT and Emerging Technologies
Neset Hikmet, University of South Carolina | nhikmet@cec.sc.edu

Decision-Making: Public Administration and Policy
Ravi Chinta, University of the District of Columbia | ravi.chinta@udc.edu

DSS, Machine Learning and Artificial Intelligence
Jin Fang, Clark University | jinfang@clarku.edu

Education, Curriculum and Cases
Ashok Viswanathan, Rutgers University | iamashokv@gmail.com
Chirag Surti, Rider University | csurti@rider.edu

Healthcare Analytics and Services Management
Dinesh Pai, Penn State Harrisburg | drp18@psu.edu

Human-Technology Interface
Joy Field, Boston College | filedjo@bc.edu

Innovation and Creativity
Afrooz Moatari-Kazerouni, Widener University | amoatarikazerouni@widener.edu

Legal, Ethical and Social Issues
Amin Keramati, Widener University | akeramati@widener.edu

Marketing and Consumer Behavior
Anshu Arora, University of the District of Columbia | anshu.arora@udc.edu

Operations Management/Operations Research
Gang Li, Bentley University | gli@bentley.edu
Neil Desnoyers, St. Joseph's University | ndesnoye@sju.edu

Strategy and Organizational Behavior
Julius Anyu, University of the District of Columbia | jndumbe@udc.edu
Sergey Ivanov, University of the District of Columbia | sivanov@udc.edu

Supply Chain Management and Logistics
Pedro Reyes, Baylor University | pedro_reyes@baylor.edu

Sustainability Management
Xiangrong Liu, Bridgewater State University | xliu@bridgew.edu

Undergraduate/Master Student Poster Competition
Jennifer Swanson, Stonehill College | jswanson@stonehill.edu
Kellyann Berube Kowalski, University of Massachusetts Dartmouth | kkowalski@umassd.edu
Kathleen Ferris-Costa, Bridgewater State University | kferriscosta@bridgew.edu
Krista Hill, Babson College | khill@babson.edu

CONFERENCE REVIEWERS

Mohammad Alshibli	State University of New York, Farmingdale State College
Ed Arnheiter	Drexel University
Amit Arora	University of the District of Columbia
Katherine Ashley	Northeastern University
Hatice ATILGAN	Eskisehir Osmangazi University
Carles Sola Belda	College of Staten Island, CUNY
Ellen Belitzky	University of New Haven
Theologos Bonitsis	New Jersey Institute of Technology
Benjamin Branch	QSBG, SEEDS Institute
Qiang Bu	Penn State
Eduardo Bugallo	UniRitter
Erin Percival Carter	University of Maine
Aichih (Jasmine) Chang	New Jersey Institute of Technology
Jimmy Chen	Bucknell University
Yucheng Chen	Commonwealth University of PA
JA-SHEN CHEN	jchen@saturn.yzu.edu.tw
Zhuowen Chen	Worcester Polytechnic Institute
Ravi Chinta	University of the District of Columbia
Q.B. Chung	Villanova University
Donovan Collier	University of the District of Columbia
Michael Craven	Marist College
Qizhi Dai	Drexel University
Jonathan Daigle	Monmouth University
Mehdi Darban	Louisiana State University - Shreveport
Kim Deranek	Nova Southeastern University
Neil Desnoyers	Saint Joseph's University
Xiaotong (Claire) Ding	Boston University
Chamila Kalpani Dissanayake	Pennsylvania State University
Sajad Ebrahimi	Wagner College
Nesreen El-Rayes	New Jersey Institute of Technology
Jin Fang	Clark University
Farbod Farhadi	Roger Williams University
Joy Field	Boston College
Charles-Henri Fredouet	ESLI School of Logistics
Yue Gao	Clark University
David Gefen	Drexel University
Tom Geurts	Bucknell University
Aarushi Gupta	Rowan University
Sarah Hertz	Empire State College, SUNY
Krista Hill	Babson College
Jae-Dong Hong	South Carolina State University
Chieh-Chih Hou	Ming Chuan University, MCU
Jian Hua	University of the District of Columbia
Samar Issa	Saint Peters University
Sergey Ivanov	University of the District of Columbia
Haleh Karimi	Bellarmine University
Atiyeh Kazeroonimonfared	Toronto Metropolitan University
Amin Keramati A	ssistant Professor, School of Business Administration, Widener University
Mehmet Kilinc	Le Moyne College
Rae Yule Kim	Montclair State University
Se Kyu Kim	Department of Sustainability Management-Inha University
SANGAHN KIM	Siena College

CONFERENCE REVIEWERS

Elif Kongar	University of New Haven
Kellyann Kowalski	University of Massachusetts, Dartmouth
Barbara Lewis	The Koppelman School of Business
Liaoliao Li	Kutztown University
Tung-Shan Liao	Management College, Yuan Ze University
Linda Boardman Liu	Boston College
Xiangrong Liu	Bridgewater State University
Chin-Yen Alice Liu	Texas A&M University – San Antonio
Kara Liu	University of Rhode Island
John Maleyeff	Boston University
Khuliso Mapila	University of Johannesburg
Andrew McInnis	State University of New York, Farmingdale State College
Alexander Mechitov	University of Montevallo
Allison Miller Georgia	College & State University
Fouad Mirzaei	University of North Texas
Afrooz Moatari-Kazerouni	Assistant Professor, School of Business Administration, Widener University
Hien Nguyen	Pennsylvania State University
Ahmet Ozkul	University of New Haven
Dinesh Pai	Pennsylvania State University
Fatma Pakdil	Eastern Connecticut State University
Arsalan Paleshi	York College of Pennsylvania
Youqin Pan	Salem state university
Leeparani Parlapalli	University of New Haven
Andrew Perry	Springfield College
Yanni Ping	St. John's University
Christine Pitocco	Stony Brook University
Yolandra Plummer	University of the District of Columbia
Loreen Powell	Commonwealth University of PA
Carmen Quirvan	The University of Rhode Island
Balaraman Rajan	California State University
Mysore Ramaswamy	Southern University and A&M College
Bhawesh sah	Northeastern University
Eda Sanchez-Persampieri	LIM College
Jeffrey Schaller	Eastern Connecticut State University
Arnold Schneider	Georgia Institute of Technology
Jessica Schwartz	University of the Cumberlands
Tamara Schwartz	York College of Pennsylvania
Rose Sebastianelli	University of Scranton
Zhengzhong Shi	University of Massachusetts, Dartmouth
Heechang Shin	Iona University
Bhupinder Sran	DeVry University
Chirag Surti	Rider University
Jennifer Swanson	Stonehill
Saeed Tabar	Assistant professor
Tih Koon Tan	University of the District of Columbia
Tulin Ural	Yeditepe University
Eddy Vertil	Rutgers University
John Visich	Bryant University
Carolyn Wang	New Jersey Institute of Technology
John Weber	DeVry University
Mei-Hsin Wu	Management College, Yuan Ze University
Gawon Yun	Missouri State University

CONFERENCE REVIEWERS

Ozdogan Zeliha	Pennsylvania State University
Jintao Zhang	Drexel University
Dongli Zhang	Fordham University
Chunjie Zhao	Clark University
Cathy Zheng	Bryant University
Haonan Zhong	

NEDSI SERVICE RECOGNITION TO PAST PROGRAM CHAIRS

Theologos Homer Bonitsis
New Jersey Institute of Technology
Martin Tuchman School of Management
Program Chair, NEDSI 2022

BOSTON COLLEGE
Carroll School of Management

NEDSI SERVICE RECOGNITION TO PAST PRESIDENTS

Gang Li
Bentley University
School of Business
President NEDSI, 2021-2022

Effie Stavrulaki
Bentley University
School of Business
President NEDSI, 2021-2022

PENN STATE HARRISBURG

We Are Penn State in the Capital Region

We'll Set You Up For Success

PennState Harrisburg

School of
Business Administration

777 West Harrisburg Pike, Middletown, PA 17057

717-948-6139 • akj11@psu.edu
harrisburg.psu.edu/business-administration

BEST PAPER AWARDS COMMITTEE

Joy Field (Chair)	Boston College
Theologos Homer Bonitsis	New Jersey Institute of technology
Neil Desnoyers	St. Joseph's University
Douglas N. Hales	University of Rhode Island
EunSu Lee	New Jersey City University
Gang Li	Bentley University
Linda Boardman Liu	Boston College
Asil Oztekin	UMass Lowell
Dinesh R. Pai	Penn State Harrisburg
Pedro M. Reyes	Baylor University
Effie Stavrulaki	Bentley Univesrity
Oya Tukel	New Jersey Institute of Technology

UNDERGRADUATE/MASTER STUDENTPOSTER COMPETITION COMMITTEE

Jennifer A. Swanson (Chair)	Stonehill College
Kathleen Ferris-Costa	Bridgewater State University
Krista Hill	Babson College
Kellyann Berube Kowalski	UMass Dartmouth

BEST PAPER AWARDS

Bryant University Best Paper Award in Supply Chain Management and Logistics

Understanding Covid-19 Impact on Supply Chain through Social Media Discussion

Suhong Li, Bryant University

Fang Chen, University of New Haven

Thomas Ngniatedema, Morgan State University

Best PhD Student Paper Award

An examination of mobile payment adoption for Taiwanese consumers: A case of the retail industry

Yong-Cheng Chang, Yuan Ze University

Best Contribution to Theory Award

Attention Based Dynamic Graph Neural Network for Asset Pricing

Ajim Uddin, New Jersey Institute of Technology

Xinyuan Tao, New Jersey Institute of Technology

Dantong Yu, New Jersey Institute of Technology

Best Application of Theory Award

I*mpacts of ESG on Gen Z and Gen Y Employees' Prosocial Motivation in the Workplace*

C. Christopher Lee, Central Connecticut State University

Heechang Shin, Iona University

Best Overall Conference Paper Award

No Treatment Without a Diagnosis: Government Regulatory Exclusion Influences on Employees Work Environments, their Behaviors, Attitude and Engagement and Perceived Performance in a Healthcare Setting

Jaime Nieto, Drexel University

David Gefen, Drexel University

UNDERGRADUATE/GRADUATE RESEARCH POSTER PARTICIPANTS

GRADUATE STUDENTS RESEARCH POSTERS

Analysis of Coronary Heart Disease
Mr. Yanjing Wang, Clark University
Mr. Ziru Wang, Clark University
Ms. Yihan Wang, Clark University

Economic Recovery in the USA after the COVID-19 Pandemic
Ms. Aakriti K C, University of New Haven
Dr. Ellen Belitzky, University of New Haven

E-Waste Analysis for India
Ms. Roopal Mishra, University of New Haven
Mr. Prittam Kumar Mahapatra, University of New Haven

Improving the Social Determinants of Health for Medicaid Recipients: A Prototype for MOOC Design
Ms. Crystal Wu, California State University, Los Angeles
Dr. Pamella Howell, California State University, Los Angeles

Integrating Efficient DOE and Intuitive Classification Trees
Ms. Jiaoxue Liu, Boston University

Predictive Modeling to Forecast Climate Change Impacts on Hospital Supply Chains in Boston
Ms. Yujue Tan, Boston University

Simulation of MBTA Repair Inventory Policies using Python
Ms. Jia Fang, Boston University

The Importance of Predictive Analytics in Healthcare Management at the Onset and During Covid-19 Pandemic
Ms. Anshu Subedi, University of New Haven
Dr. Ellen Belitzky, University of New Haven

Utilizing Game Theory to Optimize Decisions for Climate Change Uncertainties in Coastal Communities
Ms. Karla Dimitri, University of the District of Columbia
Dr. Bryan Higgs, University of the District of Columbia

UNDERGRADUATE STUDENTS RESEARCH POSTERS

Communication Technology in the Food Supply Chain
Mr. Rex Bingham, College of Charleston
Dr. Rafael Teixeira, College of Charleston

Disruptive Healthcare Technology Solutions: How Do They Improve the Healthcare Value Chain?
Ms. Shea'lyn Hubbs, Widener University
Prof. Afrooz Moatari-Kazerouni, Widener University

UNDERGRADUATE/GRADUATE RESEARCH POSTER PARTICIPANTS

Enhancing Sustainability in Supply Chain Logistics
Mr. Eduardo Bugay, University of the District of Columbia
Dr. Amit Arora, University of the District of Columbia
Dr. Anshu Arora, University of the District of Columbia

Evaluating the Magnitude of Ethical Behavior: Its Determinants and Process
Mr. Joseph Kim, New York University
Mr. Joshua Kim, Livingston High School
Mr. Philip Bae, Summit Edu Academy

Examining Human-Robot Interaction through Storytelling and Drama
Ms. Kayleah Shelton, University of the District of Columbia
Dr. Amit Arora, University of the District of Columbia
Dr. Anshu Arora, University of the District of Columbia

From Big Screen to Small Screen: Exploring the Key Differences Between Film and Social Media Product Placements and their Effects on Children's Perception
Ms. Carrie Ebbin, Brooklyn College

Generational Humor and Marketing: Appealing to the Age of Social Media
Ms. Alexa Whyte, Macaulay Honors College at Brooklyn College

How Can Marketers Leverage Artificial Intelligence for Consumers?
Ms. Camaren Rogers, University of the District of Columbia
Dr. Amit Arora, University of the District of Columbia
Dr. Anshu Arora, University of the District of Columbia

How Does Human-Robot Interaction Affect Learning and Trust Building?
Ms. Micah Hamilton, University of the District of Columbia
Dr. Amit Arora, University of the District of Columbia
Dr. Anshu Arora, University of the District of Columbia

Image Classification of Stroke Blood Clot Origin By Machine Learning Approaches
Mrs. Lily Sica, Pennsylvania State University
Mr. Alex Li, Pennsylvania State University
Dr. Hien Nguyen, Pennsylvania State University

Investigating Students' Perspectives on their College Using Exploratory Data Analysis
Ms. Joanna Smith, Valdosta State University
Mr. Andrew King, Valdosta State University
Ms. Gerniya Flewellen, Valdosta State University

Man vs. Machine: Cognitive and Artificial Intelligence
Ms. Arlene Marshall, University of the District of Columbia
Dr. Amit Arora, University of the District of Columbia
Dr. Anshu Arora, University of the District of Columbia

No Phishing Allowed: Analysis of Phishing Cybersecurity Awareness at Shippensburg University
Mr. Casey Platts, Shippensburg University
Dr. Joseph Catanio, Shippensburg University

UNDERGRADUATE/GRADUATE RESEARCH POSTER PARTICIPANTS

Over Specialization Issues within Recommender Systems
Ms. Glynnis Tan, Kean University
Ms. Yesica Espinoza-Xitumul, Kean University

Pride and Shame Appeals in Sustainable Fashion Consumption Advertisement
Ms. Cristina Heilbron, Babson College
Dr. Krista Hill, Babson College

Rowan After Hours Employment Operations
Ms. Samantha Gross, Rowan University
Ms. Grace Gerard, Rowan University
Ms. Aarushi Gupta, Rowan University

Russian Ukraine Crisis: Environmental Score Impact on Equity Returns
Mr. Nicholas Yalch, Monmouth University
Dr. Jonathan Daigle, Monmouth University

Sand Loss and Overconsumption: The Loss of Coasts and Lives
Ms. California Muratore, Bridgewater State University

The Danger of Plastic Waste
Ms. Amelia Tieri, Bridgewater State University

The Development of Social Entrepreneurship: From the Perspective of Michael Porter's Diamond Model
Mr. Joseph Kim, New York University
Mr. Philip Bae, Summit Edu Academy

The Filipino Business Mindset: Western, Eastern, or Synergic?
Ms. Glynnis Tan, Kean University
Dr. Weichun Zhu, Kean University

The Information Source Terrain: Mapping Relevancy in Supply Chain Media
Mr. Sebastiano Ronchi, College of Charleston
Dr. Rafael Teixeira, College of Charleston

The Impact of Carbon Emission Trading in Redesigning American and Asian Global Supply Chain Network for Costs and Material-based GreenHouse Gas Volumes
Ms. Miyu Kotegawa, The University of Electro-Communications
Mr. Takaki Nagao, The University of Electro-Communications
Prof. Tetsuo Yamada, The University of Electro-Communications

UNIVERSITY OF MARYLAND 18 56
ROBERT H. SMITH SCHOOL OF BUSINESS
CENTER FOR GLOBAL BUSINESS

The URI College of Business is a community of big thinkers interested in new ways of looking at the world of today—and tomorrow. Here, you will learn to solve problems, find innovative solutions, and think globally.

Our graduate programs are taught by distinguished faculty and experienced practitioners and offer a comprehensive curriculum that equips you with the critical knowledge and skills to succeed.

Join us!

THE UNIVERSITY OF RHODE ISLAND
100
1923–2023
COLLEGE OF BUSINESS

uri.edu/business

Graduate Programs

Master of Business Administration
M.S. Accounting
M.S. Textiles, Fashion Merchandising and Design
M.S. Labor Relations and Human Resources
M.S. Healthcare Management (online)
M.S. Supply Chain Management (online)
Executive Doctorate of Business Administration
Ph.D. of Business Administration

NEDSI

NORTHEAST DECISION SCIENCES INSTITUTE

Conference Program
Schedule of Presentations

March 30 – April 1, 2023

Crystal Gateway Marriott
Washington, DC Metro Area

Thursday, 30 March

7:30am **Registration**
Foyer (Lobby Level)

7:30am **Breakfast (sponsored by Penn State Harrisburg)**
Sky View (First Floor)

8am **AFE-1: Financial Asset Pricing**
Lee (Lobby Level)
Chaired by: Theologos Homer Bonitsis

Extreme Bitcoin Returns and Investors Inattention to the Stock Market: Evidence from Stock Return Co-movement, Before and During COVID-19

»Dr. Tao Li [1],Dr. Wenxiu Nan [2](1. SUNY New Paltz, 2. Salisbury University)

Examining "Boilerplate" CAM Reporting with NLP

»Dr. Nathan Slavin [1],Dr. Jianing Fang [2],Mr. David Fisher [3](1. Hofstra University, 2. Kean University, 3. USC)

Attention Based Dynamic Graph Neural Network for Asset Pricing

»Dr. Ajim Uddin [1],Dr. Xinyuan Tao [1],Dr. Dantong Yu [1](1. New Jersey Institute of Technology)

8am **BAK-1: Evaluation and Performance Measure**
Jefferson (Lobby Level)
Chaired by: Dr. Anil Aggrawal

Using learning analytics to improve students' performance

»Dr. Yougin Pan [1],Prof. Jian Gu [1](1. Salem state university)

A STRATIFICATION-BASED CLUSTERING ANALYSIS APPROACH TO PERFORMANCE EVALUATION OF DECISION-MAKING UNITS

»Prof. Jae-Dong Hong [1](1. South Carolina State University)

Using Predictive Analytics to Forecast Player Performance in Poker Tournaments

»Dr. Michael Paz [1],Dr. Mikhail Sher [2],Dr. Robert Scott III [2](1. Cornell University, 2. Monmouth University)

The Design and Operation of Digital Platforms under Folk Theories of Sociotechnical Systems

»Prof. Jordan W. Suchow [1](1. School of Business, Stevens Institute of Technology)

8am **Workshop: A SEEDS' Secure Data Repository OKD SNO Edge Framework**
Jackson (Lobby Level)

A SEEDS' Secure Data Repository OKD SNO Edge Framework

»Dr. Benjamin Branch [1],Mr. Luigi Zuccarelli [2],Mr. Thomas Glenn [3],Mr. Micheal Sullivan [3],Dr. Neset Hikmet [4],Mr. Bill Wright [5],Mr. Jarvis Green [6],Dr. Charles Zelek [7](1. QSBG, SEEDS Institute, 2. Red Hat, Principal Software Engineer, 3. SEEDS Institute, Global Supply Chain Security SME, 4. University of South Carolina, Professor of Integrated Information Technology, 5. Red Hat, Head of Edge and AI; Enterprise Neurosystem, Chairperson, 6. Green Reef Foundation, Oceans 97, Commercial Fisheries and Food Distribution SME, 7. CEO, Climate Smart Business Solutions, LLC, Global decarbonization and energy SME)

8am **DMA-1: Decision Making 1**
Madison (Lobby Level)
Chaired by: Dr. Ravi Chinta

A Decision Model for Contemporary Urban Mobility Design

»Mr. Tristan Stull [1](1. University of Massachusetts Boston)

Continued from **Thursday, 30 March**

A Mixed Approach for Highway-Rail Garde Crossings Risk Analysis Considering Crash Likelihood and Severity

»Dr. Amin Keramati [1],Dr. Pan Lu [2],Prof. Afrooz Moatari-Kazerouni [1](1. Assistant Professor, School of Business Adminstration, Widener University, 2. Associate Professor, North Dakota State University)

Application of Qualitative Judgments Based on Ordinal Model in Public Policy Decision Making

»Prof. Alexander Mechitov [1],Prof. Helen Moshkovich [1](1. University of Montevallo)

8am **DMP-1: Machine Learning & Supply Chain**

Pentagon B (First Floor)

Chaired by: Ms. Lanqing Du

Visualizing DOE Results with Classification Trees: A Supply Chain Application

»Ms. Jiaoxue Liu [1],Prof. John Maleyeff [1],Mr. Adrian Perez [1](1. Boston University)

Machine Learning in Supply Chain Management

»Mrs. Toyin Asonibare [1],Dr. Abirami Radhakrishnan [2](1. Morgan State, 2. Morgan State University)

A DEA-Based DSS for Food Shoppers

»Dr. Christine Pitocco [1],Dr. Thomas Sexton [1](1. Stony Brook University)

How Blockchain PaaS and 3PLs can help disrupted logistics network

»Ms. Lanqing Du [1],Dr. Jinwook Lee [1](1. Drexel University)

9:30am **Workshop: Integrating Immersive Technologies into the Business School Curriculum**

Lee (Lobby Level)

Integrating immersive technologies into the Business School curriculum

»Dr. Yolandra Plummer [1](1. University of the District of Columbia)

9:30am **HAS-1: What's Behind Hospital Cost Growth and Its Financial Vulnerability**

Jefferson (Lobby Level)

Chaired by: Dr. Dinesh Pai

High Cost Procedures in US Hospitals: Study of Influencing Factors

»Dr. Ravi Chinta [1],Dr. Mohamad Sepehri [1](1. University of the District of Columbia)

The Non-Clinical Determinants of Variance in Hospital Charges: Evidence from Alcohol and Drug Abuse Cases in 2019 in US (DRG = 897)

»Dr. April Nguyen [1],Dr. Ravi Chinta [1],Dr. Suzan Abed [1](1. University of the District of Columbia)

Kidney Transplant Cases in US: Study of Determinants of Variance in Hospital Charges

»Dr. Aigbe Akhigbe [1],Dr. Ravi Chinta [1](1. University of the District of Columbia)

Critical Access Hospitals and Their Financial Vulnerability

»Dr. Chamila Kalpani Dissanayake [1],Dr. Dinesh Pai [1](1. Pennsylvania State University)

Continued from **Thursday, 30 March**

IMPROVING DISPARITIES IN PROSTATE CANCER WITH TECHNOLOGY IN THE HISPANIC COMMUNITY

»Ms. Crystal Wu [1],Dr. Pamella Howell [1],Mr. Eddie Castanon [1](1. California State University, Los Angeles)

9:30am **Workshop: Adding to Your Online Teaching Toolbox**
Jackson (Lobby Level)

Adding to Your Online Teaching Toolbox

»Dr. Kathleen Ferris-Costa [1],Dr. Krista Hill [2](1. Bridgewater State University, 2. Babson College)

9:30am **IAC-1: Information Technology for Creativity**
Madison (Lobby Level)
Chaired by: Dr. Barbara Lewis

Innovation Network Characteristics: IT Value Co-Creation and Network Openness Propensity

»Dr. Mehdi Darban [1](1. Louisiana State University - Shreveport)

Fostering Employee and Organizational Resilience in the Digital Age: Can Higher Education Help?

»Dr. Barbara Lewis [1],Dr. Hershey Friedman [1](1. The Koppelman School of Business)

9:30am **LES-1: Legal, Ethical, and Social Issues**
Pentagon B (First Floor)
Chaired by: Dr. Neset Hikmet

Reciprocality of Relationships Helps Increase Resilience: Strengthening Social Support through Social-Behavior Ceuticals

»Dr. Neset Hikmet [1],Dr. Hatice ATILGAN [2](1. University of South Carolina, 2. Eskisehir Osmangazi University)

Reframing and Restructuring Organizational Strategies for Addressing Harassment and Discrimination in the Workplace

»Dr. Erin Percival Carter [1],Dr. William Obenauer [1](1. University of Maine)

Built-In Bias: The Ethical Implications of Bias in Artificial Intelligence

»Dr. Rhoda Joseph [1],Dr. Mohammad Ali [1](1. Penn State Harrisburg)

10:45am **Coffee Break (sponsored by Boston College)**
Sky View (First Floor)

11:15am **OMR-1: Decision Methods**
Lee (Lobby Level)
Chaired by: Prof. Jae-Dong Hong

A Fuzzy Linear Physical Programming Approach for Performance Evaluation and Benchmarking

»Dr. Gazi Duman [1],Dr. Elif Kongar [1],Prof. Surendra M. Gupta [2](1. University of New Haven, 2. Northeastern University)

Competitive Balance in Major League Baseball and the National Basketball Association: A Comparison of the Two Leagues

»Dr. Herbert Lewis [1],Mr. Amal Lukose [1](1. Stony Brook University)

Continued from **Thursday, 30 March**

EFFICIENCY-BASED FLEET LOCATION-ALLOCATION-ROUTING SCHEME DESIGN UNDER THE RISK OF ROUTE DISRUPTIONS

»Prof. Jae-Dong Hong [1],Dr. Judith Mwakalonge [1],Dr. Saidi Siuhi [1](1. South Carolina State University)

Cross-Channel Policies in Omnichannel Operations

»Mr. Tao Xu [1],Dr. Wenjing Shen [1](1. Drexel University)

11:15am **Panel: Role of Government in Business Resiliency**
Jefferson (Lobby Level)
Chaired by: Dr. Victor McCrary

11:15am **ECC-2: Future of Teaching - 2**
Jackson (Lobby Level)
Chaired by: Dr. Chin-Yen Alice Liu

Teaching Business Statistics in a 2-Hour Lecture

»Dr. Jinchang Wang [1](1. Stockton University)

An Investigation on the Effects of the Pandemic on Students' Attitudes Toward Technology Careers

»Dr. Kenneth Sousa [1](1. Bryant University)

Advancing Mathematical Modeling in K-12 Education

»Mr. Neil Desnoyers [1](1. Saint Joseph's University)

Visual Business Intelligence implementation in the Curriculum

»Dr. Anil Aggrawal [1],Dr. Cong Zhang [1](1. University of Baltimore)

11:15am **SCM-1: Industry 4.0 technologies**
Madison (Lobby Level)
Chaired by: Dr. Pedro Reyes

Analyzing critical success factors for drone logistics using machine learning models built on social media data

»Dr. Bhawesh sah [1],Dr. Peter Tarasewich [1],Prof. Yang Lee [1](1. Northeastern University)

Exploring the potential of supply chain business models in the Industry 4.0 era

»Dr. Pedro Reyes [1],Dr. John Visich [2],Dr. Patrick Jaska [3],Dr. Michael Gravier [2](1. Baylor University, 2. Bryant University, 3. University of Mary Hardin-Baylor)

Simulation Analysis of Repair Kits with Customized Part Service Levels

»Prof. John Maleyeff [1],Ms. Jia Fang [1],Mr. Jianbo Jiang [1](1. Boston University)

Creating value in supply network maps – a hierarchical approach

»Dr. Penina Orenstein [1],Ms. Caroline Weeks [2](1. Seton Hall University, 2. GBH (Financial Analyst))

11:15am **MCB-1: Marketing Modernism - ChatGPT, Gen Z, and Customer Choice Approach**
Pentagon B (First Floor)
Chaired by: Dr. Anshu Arora

Would Gen Z enjoy virtual tourism experience? The Effects of Virtual Tourism Characteristics on Behavioral Intention

»Prof. Chieh-Chih Hou [1],Prof. JA-SHEN CHEN [2](1. Ming Chuan University, MCU, 2. jchen@saturn.yzu.edu.tw)

Continued from **Thursday, 30 March**

A customer choice approach to retail merchandising: Evidence from the apparel industry

»Dr. Katherine Ashley [1](1. Northeastern University)

I Don't Know What That Is: The Effects of Brand Incongruence on Consumer Choices

»Dr. Rusty Stough [1](1. University of Maine)

12:30pm **Lunch Break (on your own)**

2pm **AFE-2: Studies in Accounting**
Lee (Lobby Level)
Chaired by: Dr. Nathan Slavin

Measure Annual Report Narratives Disclosure: Empirical Evidence from Disclosure of COVID-19 Information in Jordan

»Dr. Suzan Abed [1],Prof. Tih Koon Tan [1],Dr. Batao Chen [1](1. University of the District of Columbia)

History of the Classified Balance Sheets in Turkey by 1982

»Dr. Recep Pekdemir [1],Dr. Mehmet Kocakulah [1],Dr. ÜMMÜHAN ASLAN [2],Dr. Melis Ercan [3](1. University of Wisconsin La Crosse, 2. Bilecik Seyh Edebali University, 3. Istanbul University)

Supplies' Upstreamness and Investment Inefficiency: The Effect of Auditors from Downstream

»Prof. SUNGSOO KIM [1](1. Rutgers University - Camden)

How to Measure the Financial Impact of ISO 14001 Certification: Introducing New Financial Metrics

»Dr. Tom Geurts [1],Dr. Beate Klingenberg [2](1. Bucknell University, 2. FOM Hochschule für Ökonomie und Management)

2pm **BAK-2: Social Media and Marketing Analytics**
Jefferson (Lobby Level)
Chaired by: Dr. Rae Yule Kim

Mapping of Multi-Year Decision Making Strategies for Climate Change Uncertainties

»Dr. Bryan Higgs [1],Ms. Karla Dimitri [1](1. University of the District of Columbia)

Sequential Pattern Analysis Under Brand Loyalty Context

»Dr. Jin Fang [1],Dr. Hanxi Sun [2],Dr. Junhee Kim [3](1. Clark University, 2. Purdue University, 3. California State University Stanislaus)

Digital Marketing Conversion on Social Media

»Dr. Rae Yule Kim [1](1. Montclair State University)

2pm **CIT-2: Cyber Security, IT 2**
Jackson (Lobby Level)
Chaired by: Dr. Leo Hong

Industry 4.0 Technologies in Maritime Industry: Systematic Literature Review and Future Research Direction for ESG Investing

»Dr. Leo Hong [1],Dr. Se Kyu Kim [2],Ms. Kara Liu [3],Dr. Douglas Hales [3](1. Millersville University, 2. Department of Sustainability Management-Inha University, 3. University of Rhode Island)

Continued from **Thursday, 30 March**

A SEEDS' Secure Federated Framework for Global Food Security and Climate Intelligence Data Repositories

»Dr. Benjamin Branch [1],Mr. Thomas Glenn [2],Mr. Micheal Sullivan [2],Mr. Jarvis Green [3],Mr. Bill Wright [4],Dr. Neset Hikmet [5],Mr. Luigi Zuccarelli [6],Dr. Charles Zelek [7](1. QSBG, SEEDS Institute, 2. SEEDS Institute, Global Supply Chain Security SME, 3. Green Reef Foundation, Oceans 97, Commercial Fisheries and Food Distribution SME, 4. Red Hat, Head of Edge and AI; Enterprise Neurosystem Lead, 5. University of South Carolina, Professor of Integrated Information Technology, 6. Red Hat, Principal Software Engineer, 7. CEO, Climate Smart Business Solutions, LLC, Global decarbonization and energy SME)

2pm **DMA-2: Decision Making 2**
Madison (Lobby Level)
Chaired by: Dr. Ravi Chinta

The Role of Online Influencers and Data Quality in the Customer Decision-making Process

»Dr. Saeed Tabar [1](1. Assistant professor, Miller College of Business, Ball State University)

COVID-19 Vaccine Line Jumpers and Social Welfare

»Dr. Jimmy Chen [1](1. Bucknell University)

Air Connectivity Policy and Socioeconomic Factors: Evidence from Sub Saharan Africa

»Prof. Soheil Sibdari [1](1. University of Massachusetts, Dartmouth)

ORDER QUALIFIERS AND ORDER WINNERS IN URBAN DEVELOPMENT

»Dr. Yasamin Salmani [1],Dr. Amin Ariannezhad [2],Dr. Golshan Madraki [3](1. Bryant University, 2. University of Arizona, 3. Clarkson University)

2pm **DMP-2: Artificial Intelligence**
Pentagon B (First Floor)
Chaired by: Mr. Waldemar Rodriguez

THE ROLE OF ARTIFICIAL INTELLIGENCE IN ADULT LUNG CANCER DETECTION

»Mrs. Sukaynah Al Haji [1],Dr. Roderick Lee [1](1. Penn State Harrisburg)

SENTIMENT, ATTITUDE, AND INTENTION TOWARD USING ARTIFICIAL INTELLIGENCE TECHNOLOGY

»Dr. Sung Shim [1](1. Seton Hall University)

ADVANCEMENTS IN COMPUTATIONAL RHETORIC: A SURVEY

»Mr. Waldemar Rodriguez [1],Dr. Q. B. Chung [1](1. Villanova University)

3:15pm **Student Poster Competition**
Sky View (First Floor)
Chaired by: Dr. Jennifer Swanson

3:15pm **Coffee Break (sponsored by University of Maryland Center for Global Business)**
Sky View (First Floor)

3:45pm **Workshop-Effective Career Management: Balancing Focus and Variety for Research Impact**
Lee (Lobby Level)

Effective Career Management: Balancing Focus and Variety for Research Impact

»Dr. K. Sivakumar [1](1. Lehigh University)

3:45pm **HAS-2: Role of Analytics in Improving Healthcare Delivery**
Jefferson (Lobby Level)
Chaired by: Dr. Yucheng Chen

Continued from **Thursday, 30 March**

Effect of Internet Usage Behavior on the Influenza Vaccine Uptake Using Machine Learning Algorithms

»Mr. Rongxuan Wang [1],Ms. Chunjie Zhao [1],Dr. Yue Gao [1](1. Clark University)

Predicting Internet Use for Health Information: Evidence from the U.S. IPUMS Data

»Ms. Chunjie Zhao [1],Mr. Rongxuan Wang [1],Dr. Yue Gao [1](1. Clark University)

The Challenges of Healthcare Data Governance in Analytics

»Dr. Chaza Abdul -Al [1],Dr. Loreen Powell [2],Dr. Yucheng Chen [2],Ms. Gwendolyn Powell [2](1. Harrisburg University of Science and Technology, 2. Commonwealth University of PA)

Review of the Receiver Operating Characteristic (ROC) Curve Use for Medical Decision Making

»Dr. Yucheng Chen [1],Dr. Chaza Abdul -Al [2],Dr. Loreen Powell [1],Ms. Gwendolyn Powell [1](1. Commonwealth University of PA, 2. Harrisburg University of Science and Technology)

3:45pm **Workshop: Teaching Negotiations with a Procurement Simulation**
Jackson (Lobby Level)

Teaching Negotiations with a Procurement Simulation

»Dr. Chirag Surti [1],Dr. Anthony Celani [2](1. Rider University, 2. Sheridan College)

3:45pm **Workshop: Quality Analysis for Services: It's Not Like Manufacturing**
Madison (Lobby Level)

Quality Analysis for Services: It's Not Like Manufacturing

»Prof. John Maleyeff [1](1. Boston University)

3:45pm **SUM-2: Modeling in Applications of Sustainability**
Pentagon B (First Floor)
Chaired by: Dr. Eric W. Stein

Analysis of the Impact of Climate Change on Hospital Services in Boston

»Prof. John Maleyeff [1],Ms. Xiaomeng Niu [1],Ms. Yujue Tan [1],Ms. Mengya Zhao [1],Prof. David Weidman [1](1. Boston University)

Are Indoor Vertical Farms Sustainable?

»Dr. Eric W. Stein [1](1. Penn State)

Modeling Fast Fashion: Sustainability of Complementary Products

»Ms. Myan Nguyen [1],Dr. Gazi Duman [2],Dr. Elif Kongar [2](1. Westbrook High School, 2. University of New Haven)

Price Gouging at the Pumpkin Patch? Expense Neglect in Agritourism Leads to Perceptions of Price Unfairness

»Dr. Erin Percival Carter [1](1. University of Maine)

The relationship between logistics performance and environmental quality

»Dr. Yougin Pan [1](1. Salem state university)

6:30pm **NEDSI President's Reception (sponsored by New Jersey Institute of Technology). Welcome Remarks by Dr. Lawrence T. Potter, Jr., Chief Academic Officer/Provost, University of the District of Columbia**
Sky View (First Floor)

Friday, 31 March

7:30am **Registration**
Foyer (Lobby Level)

7:30am **Breakfast (sponsored by University of Rhode Island)**
Sky View (First Floor)

8am **STB-2: Strategy and Organizational Behavior 2**
Lee (Lobby Level)
Chaired by: Dr. Carles Sola Belda

The Dynamics of Paying it Forward: How and Why Leaders Transfer Past Generosity To The Future

»Dr. long wang [1](1. City University of Hong Kong)

Social preferences and framing as efficiency drivers in organizations. Experimental evidence in complex and volatile environments.

»Dr. Carles Sola Belda [1],Dr. Jessica Ellis [2](1. College of Staten Island, CUNY, 2. Booz Allen Hamilton)

To Merge or Not to Merge: Analysis of External and Internal Drivers of US Mergers and Acquisitions by Sector, 1999-2018

»Mr. Vignesh Sankaradas [1],Ms. Yue Ying [1],Dr. Kathleen Park [1],Dr. Eugene Pinsky [1](1. Boston University)

A Study on the Effect of Internal Open Innovation Activities on Digital Transformation Capability and Firm Performance: Evidence from Taiwanese SMEs

»Prof. Chieh-Chih Hou [1](1. Ming Chuan University, MCU)

8am **SCM-2: Logistics and Distribution**
Jefferson (Lobby Level)
Chaired by: Dr. Vahid Ghomi

Drone Routing Problem with Flexible Repeat Visits in Road Traffic Monitoring

»Dr. Sepideh Alavi [1](1. Assistant professor)

Determining Production and Distribution Policy in Two-Tier Supply Chain Using Reinforcement Learning

»Dr. Dmitriy Shaltayev [1],Dr. Borga Deniz [2],Mr. Joona Rahko [3](1. Christopher Newport University, 2. Framingham State University, 3. Unity Technologies)

Collaborative Logistic Vehicle Routing Problem Among Small Shippers: A New Genetic Algorithm Approach

»Dr. Vahid Ghomi [1],Dr. Sina Shokoohyar [2],Dr. Farnaz Ghazi Nezami [3],Dr. Vahid Dardashti [4](1. Penn State Mont Alto, 2. Seton Hall University, 3. Kettering University, 4. Georgia Institute of Technology)

Top Management Tournament Incentives and Inventory Efficiency

»Dr. Emma Peng [1],Dr. Dongli Zhang [1],Dr. Yuan Xie [1](1. Fordham University)

8am **MCB-2: Marketing Sensitivities and Volatilities: Cancel Culture, Human Trafficking, and Price Perceptions**
Jackson (Lobby Level)
Chaired by: Dr. Katherine Ashley

Estimating Risk of Sex and Labor Enslavement Among Consenting Human Trafficking Job Seekers: a Global Sample

»Dr. Vernon Murray [1](1. Marist College)

An examination of mobile payment adoption for Taiwanese consumers: A case of the retail industry

»Mr. Yong-Cheng Chang [1],Prof. JA-SHEN CHEN [2],Prof. Wang Ming-Chao [3](1. jeng1021@gmail.com, 2. jchen@saturn.yzu.edu.tw, 3. Management College, Yuan Ze University)

Continued from **Friday, 31 March**

Analysis of passenger satisfaction of airlines

»Mr. Yanjing Wang [1],Mr. Ziru Wang [1],Ms. Yihan Wang [1](1. Clark University)

Product Advertising and Financial Analyst Forecasts

»Dr. Seung Won Lee [1](1. Penn State Harrisburg)

8am **BAK-3: Digital Transformation, Supply Chain and Transportation**

Pentagon B (First Floor)
Chaired by: Dr. EunSu Lee

Past, Present, and Future of Digital Transformation-A Literature Review

»Mr. CHUN HUNG LIAO [1],Dr. Jonathan Ho [1](1. Management College, Yuan Ze University)

Understanding Covid-19 Impact on Supply Chain through Social Media Discussion

»Dr. Suhong Li [1],Dr. Fang Chen [2],Dr. Thomas Ngniatedema [3](1. Bryant University, 2. University of New Haven, 3. Morgan State University)

Business Intelligence Platform Development for Monitoring Bus Operation Risk Indicators

»Ms. Xiaotong (Claire) Ding [1],Prof. John Maleyeff [2],Mr. Frans Valk [1](1. Massachusetts Bay Transportation Authority, 2. Boston University)

Using GIS and network modeling for hay export

»Dr. EunSu Lee [1](1. New Jersey City University)

IS MONEYBALL EFFICIENT? A DEA Approach to Analyzing Baseball

»Dr. Brian Adams [1],Dr. Taoufik Meklachi [1](1. Penn State Harrisburg)

9:30am **BAK-4: Knowledge management and organization analytics**

Lee (Lobby Level)
Chaired by: Dr. Bharat Kolluri

Overhearing Hushed Voices: Using Unobtrusive Methods to Uncover Work Sentiments of People with Epilepsy

»Dr. Asha Rao [1],Dr. Surendra Sarnikar [1](1. California State University East Bay)

Knowledge Sharing Incentives in AI-Enabled Knowledge Management

»Dr. Shankar Sundaresan [1],Dr. Zuopeng Zhang [2](1. Rutgers University - Camden, 2. University of North Florida)

Developmental Trajectory and Trend of the Paternalistic Leadership

»Ms. Shiau-ling Wang [1],Prof. Chiung-Yi Huang [1](1. Management College, Yuan Ze University)

Exploring Expectancy Theory for Business Analysts: The Option of Failure

»Dr. Ellen Belitzky [1](1. University of New Haven)

Introduction to Business Analytics as a required core course in the undergraduate business program at the University of Hartford: Discussion on coverage of topics

»Dr. Bharat Kolluri [1],Dr. Ning Jia [1](1. University of Hartford)

Continued from **Friday, 31 March**

9:30am **DMP-3: Prediction**
Jefferson (Lobby Level)
Chaired by: Mr. Mojtaba Talaei-Khoei

A Comparative Study of EES Technologies using DEA Models

»Dr. Ahmet Akgun [1],Dr. Mehmet Yildirim [2](1. Pennsylvania Western University, 2. Wichita State University)

CorMiDEA: A Novel Feature Selection Algorithm for Improved Learning with High Dimension and Low Sample Size

»Mr. Mojtaba Talaei-Khoei [1],Dr. Asil Oztekin [1],Prof. Yao Chen [1](1. University of Massachusetts Lowell)

9:30am **Workshop: Increasing the Amount of Instructor Contact for an Engaging Online MBA Course**
Jackson (Lobby Level)

Increasing the Amount of Instructor Contact for an Engaging Online MBA Course

»Dr. Roberto Garcia [1](1. Indiana University)

10:45am **Student Poster Competition**
Sky View (First Floor)
Chaired by: Dr. Jennifer Swanson

10:45am **Coffee Break (sponsored by Georgia Tech CIBER)**
Sky View (First Floor)

11:15am **STB-3: Strategy and Organizational Behavior 3**
Jefferson (Lobby Level)
Chaired by: Dr. Sergey Ivanov

The Relationship Between the Multiple Dimensions of Commuting Satisfaction and the Appeal of Teleworking

»Dr. William Hampton-Sosa [1],Dr. Holly Chiu [1](1. The Koppelman School of Business)

Trusting the Fall: Repetitive Obstacle Performance Evaluation System (ROPES) Courses as Organization Development and Change Intervention

»Mr. Aaron Dombroski [1],Prof. Robert Yawson [2](1. Leadership and Organizational Studies, University of Southern Maine, 2. School of Business, Quinnipiac University)

Diversity and Inclusion Matter: We Are Better Together

»Dr. Sergey Ivanov [1],Ms. Bonita Harrison [1](1. University of the District of Columbia)

A Basic Leadership Inquiry – Failure of Leadership at Wells Fargo: A Case Study

»Dr. Sergey Ivanov [1],Ms. Cindy Hernandez [1](1. University of the District of Columbia)

Veteran Entrepreneurship: Review of Education Programs, Theory and Practice

»Dr. Sergey Ivanov [1],Mr. Nawras Taffal [1],Dr. Alex Maritz [2],Dr. Julius Anyu [1](1. University of the District of Columbia, 2. La Trobe University)

11:15am **SCM-3: Supply Chain Resilience**
Jackson (Lobby Level)
Chaired by: Dr. Joan Cullinane

EXPLORING THE LIMITS OF SUPPLY CHAIN ROBUSTNESS

»Prof. Henry Adobor [1](1. Quinnipiac)

Continued from **Friday, 31 March**

Developing Supply Chain Resilience to Disruption: Mindful Organizing and Supply Chain Relationships

»Dr. Joan Cullinane [1],Prof. Orakwue Arinze [1](1. Drexel University)

11:15am **Center for International Business Education and Research (CIBER) Panel**
Lee (Lobby Level)
Chaired by: Dr. John McIntyre and Dr. Roberto Garcia and Dr. Kislaya Prasad

12:30pm **NEDSI Board Meeting (by invitation only)**
Crystal Boardroom (First Floor)
Chaired by: Dr. Dinesh Pai

12:30pm **Lunch Break (on your own)**

1:30pm **NEDSI Business Meeting**
Lee (Lobby Level)

2pm **ECC-4: Analytics in Education – 2**
Lee (Lobby Level)
Chaired by: Dr. Bhawesh sah

Data Analysis of the Teaching of Data Analytics Courses

»Dr. Derald Wentzien [1](1. Delaware State University)

Analytics Skill Building: Leveraging R in the First Statistics Course

»Dr. Linda Boardman Liu [1](1. Boston College)

Shifting Forums : Establishing MS in Business Analytics Program

»Prof. Yang Lee [1],Dr. Peter Tarasewich [1],Dr. Bhawesh sah [1](1. Northeastern University)

Healthcare Analytics Teaching Tools: Cases for Ensuring Uniform Dates

»Dr. Yucheng Chen [1],Ms. Gwendolyn Powell [1],Dr. Chaza Abdul -Al [2],Dr. Loreen Powell [1](1. Commonwealth University of PA, 2. Harrisburg University of Science and Technology)

Does Management Education Need a Facelift? The Intersection of Managing, Leading, and Coaching. Part III

»Dr. Maureen Mackenzie [1](1. Molloy University)

2pm **HAS-4: Miscellaneous Topics in Healthcare Performance Management**
Jefferson (Lobby Level)
Chaired by: Dr. Neset Hikmet

What are Measurement-based Care (MBC) software service providers saying in a world where MBC can transform mental healthcare?

»Mr. Donald Jenkins [1],Dr. Josephine Namayanja [2],Ms. Rebecca Namubiru [3](1. Mirah, 2. University of Massachusetts, Boston, 3. Hochschule Neu-Ulm)

On the Role of Trust and Previous Psychological Contract Violation in Consumer Switching to Telemedicine Services

»Dr. Dodi Mossafer [1],Dr. David Gefen [2],Dr. Qizhi Dai [2](1. Accenture, 2. Drexel University)

Continued from **Friday, 31 March**

An Examination of Relationship between COMMUNICATION QUALITY, PERFORMANCE Score, AND RECOMMENDATION of A HOSPITAL using IMPORTANCE-PERFORMANCE MAP ANALYSIS

»Dr. Neset Hikmet [1],Prof. Tulin Ural [2],Dr. Emrah ATILGAN [3](1. University of South Carolina, 2. Yeditepe University, 3. Eskisehir Osmangazi University)

ISO 9001:2015 Quality Standard Implementation: Constraints or a Source of Capabilities for Healthcare Resilience?

»Dr. Bill Ritchie [1],Dr. Steven Melnyk [2](1. James Madison University, 2. Michigan State University)

A CASE PRESENTATION OF A CLIENT WITH A DIAGNOSIS OF SCHIZOPHRENIA IN A COMMUNITY CARE UNIT IN MELBOURNE, AUSTRALIA

»Mr. Qian He [1],Dr. Michael Ha [2],Dr. Tan Kan Ku [3](1. Stockton University, 2. Beijing Institute of Technology, 3. Institute of Health and Management)

2pm **OMR-4: Quality and Project Management**
Jackson (Lobby Level)
Chaired by: Dr. Ed Arnheiter

New model for measuring criticality in project management networks

»Dr. Fariborz Partovi [1](1. Drexel University)

Improving Police Response with a Crisis Intervention Team

»Ms. Patricia McGrath-Rodriguez [1],Mr. Temirbek Zhorobaev [1],Ms. Mercy Karani [1],Dr. Borinara Park [1](1. Illinois State University)

The Case for Integrating Lean Management with Pharmacy Automation

»Dr. Ed Arnheiter [1],Mr. Alex Arnheiter [2](1. Drexel University, 2. Resonetics)

2pm **SCM-4: Networks**
Madison (Lobby Level)
Chaired by: Dr. David Cortes

Effects of Consumer Showrooming and Supplier Encroachment on an Omnichannel Retailer

»Prof. Samayita Guha [1],Prof. Abhishek Roy [2],Prof. Subodha Kumar [2](1. Florida International University, 2. Fox School of Business at Temple University)

Proximity Framework: Supply Chain's Link

»Dr. David Cortes [1](1. Providence College)

The impact of supply chain disruptions and 3D printing investments

»Dr. Kraiwinee Bunyaratavej [1](1. Delaware State University)

2pm **AFE-5: Topics in Financial Valuation II**
Pentagon B (First Floor)
Chaired by: Dr. John Malindretos

A Leveraged Bootstrap Causality approach to investigate relationships among 11 GICS stock market sectors and the Covid-19 cases after the declaration of Pandemic

»Dr. Howard Lee [1](1. Bloomsburg University of PA)

Are Commercial Lending Decisions Affected by Severity of Internal Control Material Weaknesses and Tolerance for Ambiguity?

»Dr. Arnold Schneider [1](1. Georgia Institute of Technology)

Continued from **Friday, 31 March**

3:15pm **Student Poster Competition**
Sky View (First Floor)
Chaired by: Dr. Jennifer Swanson

3:15pm **Coffee Break (sponsored by Indiana University CIBER)**
Sky View (First Floor)

3:45pm **ECC-5: DEI in Education**
Lee (Lobby Level)
Chaired by: Dr. Roderick Lee

Overcoming the Barriers to Participation in Quality High Impact Practices for Structurally Disadvantaged Students in STEM

»Dr. Roderick Lee [1],Mr. Adam Lenker [1](1. Penn State Harrisburg)

AN EXAMINATION OF THE ROLE OF INCLUSIVE, EFFECTIVE, AND INNOVATIVE PEDAGOGICAL PRACTICES IN STEM EDUCATION

»Mr. Chandra Niredi [1],Dr. Roderick Lee [1](1. Penn State Harrisburg)

Inclusive Design Thinking: A Model for Architecting Inclusive STEM Classrooms for Women

»Ms. Pavani Sunku [1],Dr. Roderick Lee [1](1. Penn State Harrisburg)

The Impact of Live Lectures on Student Engagement and Performance in Online Classes

»Dr. John Weber [1],Dr. Bhupinder Sran [1](1. DeVry Univesity)

Analysis of the factors impacting enrollment in higher education institutions using Interpretive Structural Modeling

»Dr. Arsalan Paleshi [1],Dr. Aman Gupta [2](1. York College of Pennsylvania, 2. Embry-Riddle Aeronautical University Worldwide)

3:45pm **SCM-5: Closed-loop Supply Chains**
Jefferson (Lobby Level)
Chaired by: Mr. Hiromasa Ijuin

Integrating Global and Closed-Loop Supply Chains to Minimize Costs and GHG Emissions using LPP

»Mr. Hiromasa Ijuin [1],Dr. Yuki Kinoshita [1],Prof. Tetsuo Yamada [1],Prof. Surendra M. Gupta [2](1. The University of Electro-Communications, 2. Northeastern University)

Designing Reverse Supply Chain Networks with Returned Product Quality Control

»Ms. Sahar Ebrahimi Bajgani [1],Prof. Sara Saberi [2],Prof. Fuminori Toyasaki [3](1. Worcester Polytechnic Insti, 2. Worcester Polytechnic Institute, 3. York University)

Blockchain-based lithium-ion battery closed-loop supply chain safety passport

»Ms. Zhuowen Chen [1],Prof. Joseph Sarkis [1],Dr. Abdullah YILDIZBASI [1],Prof. Yan Wang [1](1. Worcester Polytechnic Institute)

3:45pm **AFE-6: International Financial Issues**
Jackson (Lobby Level)
Chaired by: Theologos Homer Bonitsis

An Examination of Operating Foreign Exchange Exposure

»Dr. Augustine Arize [1],Dr. John Barkoulas [2],Dr. John Malindretos [3],Dr. Alex Panayides [4],Dr. Chuanqian Zhang [4](1. Texas A&M University – Commerce, 2. Georgia southern, 3. william Paterson University, 4. William Paterson University)

Continued from **Friday, 31 March**

Drivers and pass-through of the EU ETS price: evidence from the power sector

»Dr. Yiyi Bai [1],Dr. Samuel Okullo [2](1. University of the District of Columbia, 2. World Bank)

Impact of Rising Diesel Prices and Truck Driver Availability on Food Transportation and Distribution

»Ms. Fuqin Zhou [1],Dr. Aichih (Jasmine) Chang [1],Dr. Jim (Junmin) Shi [1](1. New Jersey Institute of Technology)

Linking Efficiency, Profitability and Productivity Measures, with Application to the U.S. Airline Industry

»Dr. Michael Tannen [1](1. University of the District of Columbia)

Explainability of Predictors in Asset Pricing Using Machine Learning and LIME

»Mr. Bhaskar Goswami [1],Dr. Ajim Uddin [1](1. New Jersey Institute of Technology)

3:45pm **ECC-6: Education, Curriculum and Cases**
Madison (Lobby Level)
Chaired by: Prof. Pradeep K Behera

United States Postal Services

»Ms. Brittney Stargell [1],Prof. Shaoping Zhao [1](1. Stockton University)

Preparing of Industry Ready Civil Engineering Students through Transformation of Capstone Design Course at UDC

»Prof. Pradeep K Behera [1](1. University of the District of Columbia)

Nothing Cryptic About the Importance of Internal Control

»Dr. Jonathan Daigle [1],Dr. Ronny Daigle [2](1. Monmouth University, 2. Sam Houston State University)

We've got this: Creating High-Impact Online Capstone Experiences for Business Programs

»Dr. Vallari Chandna [1],Mr. Praneet Tiwari [1],Mr. Anup Nair [1](1. University of Wisconsin-Green Bay)

3:45pm **Meet the Editors Panel: Dr. K. Sivakumar (Associate Editor of Journal of Business Research), Dr. Jimmy Chen (Area Editor of Decision Sciences Journal of Innovative Education), Dr. Xenophon Koufteros (Editor-in-Chief of Decision Sciences Journal)**
Pentagon B (First Floor)

5:15pm **Keynote by Mr. Rhom Erskine, VP-Global Diversity and Inclusion, Lockheed Martin**
Grand Ballroom (First Floor)

5:45pm **Fireside Chat with Mr. Michael Kenny, VP-Operational Risk at Freddie Mac; Mr. Thadi Murali, Managing Principal at Capco Consulting; and Ms. Kelly Siu, Privacy Risk Program Manager at Meta (fka Facebook)**
Grand Ballroom (First Floor)

6:30pm **Gala Awards Dinner (sponsored by University of the District of Columbia)**
Grand Ballroom (First Floor)

Saturday, 1 April

8am **ECC-1: Future of Teaching - 1**
Lee (Lobby Level)
Chaired by: Dr. Arsalan Paleshi

Has Higher Education Lost Its Way? Or Just Its Credibility?

»Dr. Hershey Friedman [1],Dr. Taiwo Amoo [1],Dr. Barbara Lewis [2](1. The koppleman school of Business, 2. The Koppelman School of Business)

Student Attitudes towards Multiple Attempts and Feedback on Quizzes

»Dr. Vitaly Brazhkin [1](1. University of West Florida)

Creating a Culture of Connectedness and Collaboration in Asynchronous Learning

»Dr. Kellyann Kowalski [1],Dr. Jennifer Swanson [2](1. University of Massachusetts, Dartmouth, 2. Stonehill College)

The integration of Collaborative University Business Experiences (CUBEs) and Collaborative Online International Learning (COIL) in Teaching

»Dr. Xiangrong Liu [1],Prof. Annett Grossmann [2](1. Bridgewater State University, 2. HHN)

8am **IAC-2: Business Innovation and Creativity**
Jefferson (Lobby Level)
Chaired by: Prof. Edward Chen

Technology Innovation for Small Businesses – Creating an online presence during a time of rapid change

»Dr. Jas Gaurav Singh [1],Dr. Elif Kongar [2](1. Nichols College, 2. University of New Haven)

OSS Project Community (OSSPC) Ambidexterity: Its Social Network Antecedents and Community Performance Impact

»Dr. Zhengzhong Shi [1](1. University of Massachusetts, Dartmouth)

Smart Farming for Sustainable Agriculture

»Prof. Edward Chen [1](1. University of Massachusetts Lowell)

8am **ECC-3: Analytics in Education – 1**
Jackson (Lobby Level)
Chaired by: Dr. Ellen Belitzky

A Knowledge Management Solution for Business Analytics Education

»Dr. Ellen Belitzky [1],Ms. Leeparani Parlapalli [1](1. University of New Haven)

Reflections on Teaching Information Visualization in Business Analytics Courses

»Dr. Peter Tarasewich [1](1. Northeastern University)

A Data Envelopment Analysis Approach to Rank Administrations of United States Academic Institutions

»Dr. Richard Muszynski III [1](1. Wilkes University)

Classroom Innovation: Teaching Undergraduate Business Analytics Using Publicly Available, Free, Real World Data

»Prof. Allison Miller [1](1. Georgia College & State Univerity)

Effective use of Teaching Intervention and Simulation Games: A Study in a Supply Chain & Operations Management Course

»Dr. Chin-Yen Alice Liu [1](1. Texas A&M University – San Antonio)

Continued from **Saturday, 1 April**

8am **HAS-3: Exploring Relationship Between Human Factors and Care Delivery**
Madison (Lobby Level)
Chaired by: Dr. Neset Hikmet

A Case Study on Scheduling and Human Resource Allocation for Surgical Procedures

»Prof. Afrooz Moatari-Kazerouni [1],Dr. Amin Keramati [1](1. Assistant Professor, School of Business Adminstration, Widener University)

Soft Skills as The Driving Force of the Fourth Industrial Revolution

»Dr. Haleh Karimi [1](1. Bellarmine Univeristy)

The Effect of Socio-Demographical Characteristics on Leader Member Interaction During the Covid-19 Pandemic Process: An Evaluation of Healthcare Professionals

»Dr. Neset Hikmet [1],Prof. Gamze YORGANCIOĞLU TARCAN [2],Prof. Meltem SAYGILI [3](1. University of South Carolina, 2. Hacettepe University, 3. Kırıkkale University)

Impact of COV-19 Pandemic on burnout and job satisfaction among emergency department health professionals

»Dr. Neset Hikmet [1],Prof. Menderes Tarcan [2](1. University of South Carolina, 2. Eskisehir Osmangazi University)

Measuring Spectator Sports Quality: Data Analytics on Co-creation of Value

»Dr. Ahmet Ozkul [1],Dr. Gazi Duman [1](1. University of New Haven)

8am **OMR-3: Operations Research Applications**
Salon G (First Floor)
Chaired by: Ms. Ashley Gekpe

Airline Ancillary Revenue Willingness to Pay during Covid 19 Pandemic

»Dr. Fouad Mirzaei [1](1. University of North Texas)

Optimum Placement of Electric Scooters to Pickup Stations

»Dr. Rajeev Kumar [1](1. Kutztown University of Pennsylvania)

Power of Community: Engaging Community to Mitigate Terrorism Risk

»Ms. Ashley Gekpe [1],Prof. Ki Young Jeong [1],Prof. Gene Shan [1](1. University of Houston Clear Lake)

9:30am **OMR-2: Operations and Inventory Management**
Lee (Lobby Level)
Chaired by: Dr. Javad Paknejad

Properties of job sequences for the three-machine flowshop problem with sequence-dependent family setups to minimize makespan

»Prof. Jeffrey Schaller [1],Dr. Jatinder Gupta [2](1. Eastern Connecticut State University, 2. University of Alabama Huntsville)

Understanding Opaque Selling from an Inventory-control Perspective

»Mr. Yuan Qu [1],Dr. Jian Yang [1](1. Rutgers University)

AN EOQ MODEL WITH PLANNED SHORTAGES AND GEOMETRICALLY DECLINING VARIABLE PROCUREMENT COST PER UNIT

»Dr. Javad Paknejad [1],Dr. John Affisco [1](1. Department of Information Systems and Business Analytics, Hofstra University)

A Study of Agile and Lean Supply Chain Design

»Prof. Gang Li [1],Prof. Yusen Xia [2](1. Bentley University, 2. Georgia State University)

Continued from **Saturday, 1 April**

9:30am **CIT-1: Cyber Security, IT 1**
Jackson (Lobby Level)
Chaired by: Dr. Bharat Kolluri

High-level vulnerabilities: The role of age in information security awareness and attitudes

»Dr. Gregory Lyon [1](1. Georgetown University)

Examining the Relationship Between Social Media and Cryptocurrencies

»Dr. Subhasish Dasgupta [1],Dr. Yuan Xue [2],Dr. Michael Chuang [3](1. George Washington Universtiy, 2. Elizabeth City State University, 3. University of Illinois)

Cultural Values and Artificial Intelligence

»Dr. Subhasish Dasgupta [1],Dr. Babita Gupta [2],Dr. Yuan Xue [3](1. George Washington Universtiy, 2. California State University Monterey Bay, 3. Elizabeth City State University)

9:30am **Workshop: Managing Social Robotics and Immersive Technologies in the Classroom**
Jefferson (Lobby Level)
Chaired by: Dr. Anshu Arora

Managing Social Robotics and Immersive Technologies in the Classroom

»Dr. Anshu Arora [1],Dr. Amit Arora [1],Mr. Timothy Gifford [2](1. University of the District of Columbia, 2. MOVIA Robotics)

11am **SUM-3: Land, Sand and Plastic Management**
Lee (Lobby Level)
Chaired by: Ms. Cassidy Buhler

Blockchain Technology for the Plastic Crisis: A Mathematical Model for Sustainable Plastic Management

»Ms. Nesreen El-Rayes [1],Dr. Aichih (Jasmine) Chang [1],Dr. Jim (Junmin) Shi [1](1. New Jersey Institute of Technology)

Optimal land conservation decisions for multiple species

»Ms. Cassidy Buhler [1],Prof. Hande Benson [1](1. Drexel University)

Impacts of ESG on Gen Z and Gen Y Employees' Prosocial Motivation in the Workplace

»Dr. C. Christopher Lee [1],Dr. Heechang Shin [2](1. Central Connecticut State University, 2. Iona University)

11am **SUM-1: ESG and Sustainability in Supply Chain**
Jefferson (Lobby Level)
Chaired by: Dr. Rose Sebastianelli

Corporate Governance and Voluntary Disclosure: The Moderating Role of Industry

»Dr. Nabil Tamimi [1],Dr. Rose Sebastianelli [1](1. University of Scranton)

THE ROLE OF SMALL BUSINESSES IN SUPPLY CHAIN SUSTAINABILITY

»Dr. Mysore Ramaswamy [1](1. Southern University and A&M College)

11am **STB-1: Strategy and Organizational Behavior 1**
Jackson (Lobby Level)
Chaired by: Dr. Jun Wu

Continued from **Saturday, 1 April**

Effect of External Financing on Sustainable Innovation: The Threshold Effect

»Dr. Xun Zhang [1],Dr. Biao Xu [2],Dr. Ning Zhang [3],Dr. Jun Wu [4](1. Hohai Univercity, 2. Nanjing University, 3. Northeastern University, 4. Georgia Gwinnett College)

No Treatment Without a Diagnosis: Government Regulatory Exclusion Influences on Employees Work Environments, their Behaviors, Attitude and Engagement and Perceived Performance in a Healthcare Setting.

»Mr. Jaime Nieto [1],Dr. David Gefen [1](1. Drexel University)

The impact of relative advantage in digital platform implementation by SMEs in South Africa

»Ms. Khuliso Mapila [1],Mr. Garth Balabanoff [2](1. University of Johannesburg, 2. University of the Witwatersrand)

11am **AFE-4: Empirical Research in Corporate Finance**
Madison (Lobby Level)
Chaired by: Dr. Recep Pekdemir

Corporate Social Responsibility Reporting Improvements and Mandated CO2 Reductions in the Northeast United States: New Evidence on Positive Outcomes from an Unexpected Symbiotic Relationship

»Dr. A J Stagliano [1](1. Saint Joseph's University)

How Do Investors Process Pay Equity Disclosures?

»Dr. Michael Craven [1](1. Marist College)

Exposing investor sentiment in stock returns

»Dr. Qiang Bu [1](1. Penn State)

Timing Value versus Growth

»Dr. Pawan Madhogarhia [1](1. York College of Pennsylvania)

11am **CIT-3: Cyber Security, IT 3**
Salon G (First Floor)
Chaired by: Dr. Neset Hikmet

CYBERCHONDRY (CSS-12) SEVERITY SCALE: A SPECIFIC APPLICATION TO HEALTH INFORMATION

»Dr. Neset Hikmet [1],Prof. Gamze YORGANCIOĞLU TARCAN [2],Prof. Meltem SAYGILI [3],Prof. Ayca Karahan [4],Mr. Nihat Baris Sebik [5](1. University of South Carolina, 2. Hacettepe University, 3. Kırıkkale University, 4. Pamukkale University, 5. Republic of Türkiye Ministry of Health)

Review of Cloud Computing Research: A Pilot Study

»Dr. Zhengzhong Shi [1](1. University of Massachusetts, Dartmouth)

The Effect of Perceived Blockchain Technology Benefits on ESG Disclosure Trustworthiness in South Korea

»Dr. Se Kyu Kim [1],Dr. Leo Hong [2],Dr. Jong Dae Kim [3](1. Department of Sustainability Management-Inha University, 2. College of Business - Millersville University, 3. College of Business Administration - Inha University)

Organizational (Digital) Response to Turbulent Business Environment due to the Pandemic

»Dr. Arvind Gudi [1],Dr. Ravi Chinta [2],Dr. Yuliya Yurova [1],Dr. Kim Deranek [1](1. Nova Southeastern University, 2. University of the District of Columbia)

PennState
Great Valley

IS PROUD TO SPONSOR THE

2023 NEDSI ANNUAL CONFERENCE

We are committed to educating current and future leaders through our high-quality, innovative, and technologically progressive opportunities for lifelong learners. We offer graduate and continuing education programs regionally in Malvern.

PROGRAMS INCLUDE:

- **Accounting**
- **Business**
- **Data Analytics**
- **Engineering**
- **Finance**
- **Leadership Development**
- **Technology**

Our locations in Malvern offers conference services which include auditorium seating, meeting rooms, and state-of-the-art technology. We work with businesses and organizations to create, host, and support customized events in an academic environment.

LEARN MORE

greatvalley.psu.edu

610-648-3248 | gvinfo@psu.edu

case centre

Discover The Case Centre: your one-stop resource for cases

The Case Centre is the independent home of the case method, dedicated to sharing knowledge, wisdom and experience to inspire and transform business education across the globe.

Free product updates

Did you know that we hold the world's largest and most diverse collection of business and management cases?

To help you stay up to date with new materials in your chosen areas of study we provide a free monthly email update service giving details of items registered during the preceding month.
www.thecasecentre.org/emailupdates

Supporting case teachers and writers

We also provide an array of services, resources and activities to support the case community. Find out about our case awards, competitions, scholarships, webinars and workshops on our website.

www.thecasecentre.org

Call for Papers
NORTHEAST DECISION SCIENCES INSTITUTE
53RD Annual Meeting,
April 4-6, 2024
Hyatt Regency,
Cambridge, Massachusetts
Submission Deadline: **December 29, 2023**
NEDSI is the Northeast Regional Subdivision of the Decision Sciences Institute

MEETING ANNOUNCEMENT

The 2024 Northeast Decision Sciences Annual Meeting will be held between **April 4-6, 2024,** in Cambridge, MA. The conference will start on Thursday and continue through Saturday night. All paid attendees are invited to the Welcome session, paper presentations, the President's Reception, Award Ceremony, and other events.

CALL FOR PAPERS

Full papers, abstracts, and workshops are invited. Undergraduate students may submit proposals for Poster Sessions. Sessions involving practitioners will be given consideration. Submissions will be refereed, and accepted papers will be published in the online Conference Proceedings. By submitting a manuscript, the author certifies that it is not copyrighted or previously published, has not been presented or accepted for presentation at another professional meeting, and is not currently under review for presentation at another professional meeting. At least one of the authors must certify that he/she intends to register for and attend the conference to present the paper if it is accepted.

All papers, abstracts, and undergraduate posters must be submitted electronically on or before December 29, 2023. If you have proposals for workshops or roundtable discussions, please email the co-chairs.

OUTSTANDING PAPER AWARDS

Outstanding papers are eligible for awards:

- Best Application of Theory
- Best Contribution to Theory
- Best Paper on Innovation or Sustainability
- Best Paper in Supply Chain Management and Logistics
- David M. Levine Award for Innovative Education
- Richard Briotta Award for Knowledge Management/Strategy
- Best Student Papers

PROGRAM TRACKS

- Accounting, Finance, Economics
- Big Data and Business Analytics
- Cyber Security, IT, IS, and Emerging Technologies
- Decision Support Systems (DSS), Machine Learning, and Artificial Intelligence
- Education, Curriculum, and Cases
- Healthcare Analytics
- Human Technology Interface
- Innovation and Creativity
- Legal, Ethical, and Social Issues
- Marketing and Consumer Behavior
- Operations Management/Operations Research
- Public Administration and Policy
- Service Management
- Strategy, Knowledge Management and Organizational Behavior
- Supply Chain Management and Logistics
- Sustainability
- Teaching and Innovative Education
- Undergraduate/Master Student Posters

HOTEL ARRANGEMENTS

The host hotel is the Hyatt Regency, overlooking Boston and the Charles River. The room reservation link will be available by June 2023. For reservations prior to then please call 866-238-4218 and ask for the NEDSI 2024 conference in April 2024.

CONFERENCE WEBSITE/SUBMISSION LINK

- The link will be available after June 2023. Check anytime at nedsi.decisionsciences.org.

CONTACT INFORMATION

2024 NEDSI Program Co-Chairs

Prof. Neil Desnoyers,
(ndesnoye@sju.edu)
Decision& System Sciences Department,
Haub School of Business,
Saint Joseph's University,
Philadelphia, PA

Dr. Xiangrong Liu
(xliu@bridgew.edu)
Department of Management and Marketing,
Ricciardi College of Business,
Bridgewater State University,
Bridgewater, MA

Thank you to our sponsors!

Table of Contents

Papers

Accounting, Finance, Economics

Big Data, Analytics, and Knowledge Management

Cyber Security, IT, and Emerging Technologies

Decision Making: Public Administration and Policy

DSS, Machine Learning, and Artificial Intelligence

Education, Curriculum, and Cases

Healthcare Analytics and Services Management

Human-Technology Interface

Innovation and Creativity

Legal, Ethical, and Social Issues

Operations Management / Operations Research

Sustainability Management

Strategy and Organizational Behvior

Supply Chain Management and Logistics

Marketing and Consumer Behavior

Abstracts

Accounting, Finance, Economics

Big Data, Analytics, and Knowledge Management

Cyber Security, IT, and Emerging Technologies

Decision Making: Public Administration and Policy

DSS, Machine Learning, and Artificial Intelligence

Education, Curriculum, and Cases

Healthcare Analytics and Services Management

Human-Technology Interface

Innovation and Creativity

Legal, Ethical, and Social Issues

Operations Management / Operations Research

Sustainability Management

Strategy and Organizational Behvior

Supply Chain Management and Logistics

Marketing and Consumer Behavior

Papers

Accounting, Finance, Economics

Attention Based Dynamic Graph Neural Network for Asset Pricing

Abstract

Recent studies suggest that networks among firms (sectors) play an essential role in asset pricing. However, it is challenging to capture and investigate the implications of networks due to the continuous evolution of networks in response to market micro and macro changes. This paper combines two state-of-the-art machine learning techniques to develop an end-to-end graph neural network model and shows its applicability in asset pricing. First, we apply the graph attention mechanism to learn dynamic network structures of the equity market over time and then use a recurrent convolutional neural network to diffuse and propagate firms' fundamental information into the learned networks. Our model is effective in both return prediction and improving portfolio performance. The result persists in different sensitivity tests and simulated data. We also show that the dynamic network learned from our model is able to capture major market events over time.

Keywords: Machine Learning, FinTech, Neural Network, Asset Pricing, Financial Network, Graph Convolutional Neural Networks

`JEL Classification:` C33, C52, C63, G10, G14

1. Introduction

Firms do not operate independently in the marketplace. They are influenced by each other through multiple channels, including but not limited to their supply chain networks, board of directors, fundamentals similarity, industry sector, and market condition. As a result, the performance of a firm depends not only on its own operation but also on other relevant firms, i.e., the interconnection among firms affects each other and their market prices. This interconnectedness among firms through contemporaneous links forms the firms'

network. Previous studies suggest that the structure, properties, and dynamism of a network provide important insights into how information flows and shocks propogate across firms and thereby affect their stock prices (Cohen & Frazzini, 2008; Muslu et al., 2014; Grullon et al., 2014; Herskovic, 2018).

Although there is a general consensus about the importance of network in asset pricing, disputes exist on the proper representation of the firms network. Sector similarity (Herskovic, 2018), correlation of historical return (Namaki et al., 2011; Tumminello et al., 2010; Di Cerbo & Taylor, 2020), cross-correlation of distress probability (Buraschi & Porchia, 2012), customer-supplier network (Herskovic et al., 2020), volatility connectedness (Diebold & Yılmaz, 2015) are used in the literature to represent financial networks.

There are two main concerns for the current representation. First, sector similarity and supply-chain network are mostly represented in a static graph and thereby fail to account for the dynamic nature of financial networks that contributes differently to returns. Although using industry sector similarity or supply-chain as the linkage between firms might reveal common information about their shared marketspace, logistic exchanges, and operational environments. As Figure 1 shows, being in the same sector or industry group does not necessarily lead to similar returns. Second, the similarity-based network predominantly uses the Pearson correlation of firms' historical returns or volatility to capture the linkage among them. The Pearson correlation only reveals the linear relationship among entities and might not be sufficient to model the nonlinear inter-dependency among stocks. In addition, these types of networks usually ignore the firm's fundamental characteristics and other latent factors that affect firms' underlying connections and their continuous evolution.

To address these two challenges, this paper proposes DY-GAP (Dynamic Graph Neural Network for Asset Pricing) – a novel two-step graph learning model to capture the dynamic nonlinear interconnections among firms and investigates their contributions to the stock price movement. In the first step, we use an attention function to learn network representation among firms at each time point. In the second step, we model spatial-temporal relationships

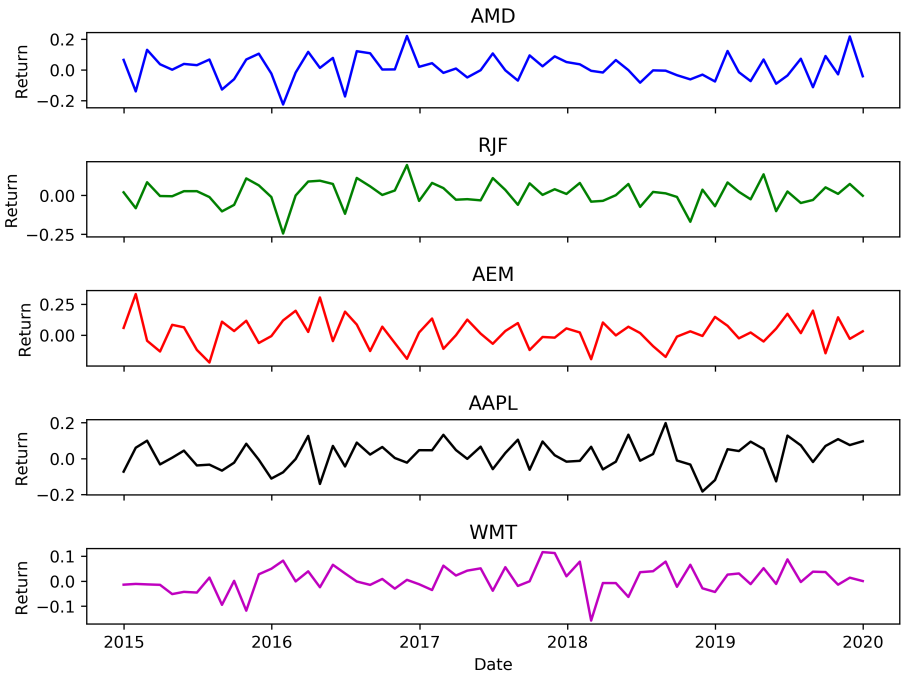


Figure 1: Rate of return from five assets over time. The timing of earning returns by AMD (Technology) is similar to that of RJF (Financial Services) and almost opposite to that from AEM (Materials-Mining). AMD and AAPL (Technology) are from the same industry sector, whereas their returns vary significantly. The dynamic nature of the changing relations among firms' return is also visible from the trend lines of AEM and WMT (Consumer Discount Stores). AEM and WMT started with little correlation in the early 2017, began to have strong co-movements between early 2017 and mid-2018, and then diverged into opposite movements from mid-2018 to December 2019.

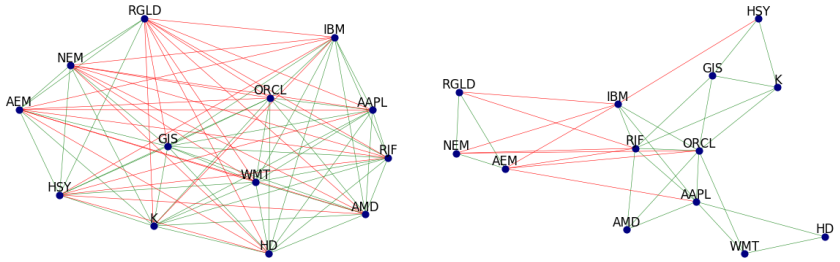
among firms by combining graph convolutional neural network and recurrent neural network to predict future returns.

Our framework is inspired by two advanced machine learning techniques for graph learning and panel data modeling. The first one is the attention mechanism. Attention plays a vital role in information processing in the brain. It is widely studied in multiple fields, including psychology and neuroscience (Borji & Itti, 2013), and recently makes its way into artificial intelligence and machine learning (Cho et al., 2015). Attention is far from clearly understood in neuroscience. Nevertheless, it is commonly accepted that human has remarkable attention capability in directing the limited cortex to various cognitive tasks. Inspired by these brain functions of arousal, alertness, and vigilance, Bahdanau et al. (2015) proposes the attention mechanism to learn data-driven latent embedding based on the relative importance of the features of each observation. Basic intuition is similar to the human cognitive

process. When asked a specific question, the machine learning model selectively concentrates on one or several things while ignoring others. Veličković et al. (2018) successfully use the attention function on graphs to assign edge coefficient to each node based on their relative importance.

We use the attention function to learn the interaction (inter-dependency) coefficient between each firm and its neighbors in a pre-selected network at each time point based on the observed features, historical return, and return correlation. Using a neural network, the attention function estimates the attention score for each firm by attending to its neighbors, given the representation of its own feature as the query. As a result, the learned network is dynamic and based on multiple aspects of firm connectedness, including their explicit and implicit interactions. This attention function also allows us to denoise the network by removing less important connections. Figure 2a shows that without any filtering approach, the Pearson correlation network creates an almost complete graph with every node connected to all other nodes in the network. In this case, the information content of the network gets intertwined with noises because all connections are treated equally important, while some linkage is actually less important. A better filtering technique can remove unnecessary edges and create a meaningful graph. In Figure 2b, the network groups firms in their respective peers and displays a better spectral location.

The second is the diffusion mechanism to model Spatio-temporal information. Empirical evidence also suggests that diffusion of common information is a major source of the lead-lag effect in stock returns (Hou, 2007; Wang & Xie, 2010). Diffusion processing in the continuous space can be extended to the discrete space (graph) in the format of random walk-in networks (Huang et al., 2015). This idea inspired a series of works in graph neural networks: Graph diffusion convolution (GDC) that incorporates diffusion mechanism to improve graph learning (Klicpera et al., 2019), node classification (Atwood & Towsley, 2016), and traffic prediction (Li et al., 2019). In this work, we follow the recurrent convolutional diffusion process proposed in Li et al. (2019) to propagate firms' information to their neighbors. The



(a) Correlation Network (b) Filtered Network

Figure 2: The network structure of a few major US firms. The levels represent the ticker symbol, and green (red) links represent the positive (negative) edge between two firms. The spectral location in Figure (b) clearly cluster firms in their industry peers. This figure also denotes the significance of negative links. The general practice of ignoring the negative signs of links will mix up different industry sectors by overlooking significant disassociation between mining firms (RGLD, NEM, AEM) and technology firms (ORCL, IBM, AAPL, AMD).

convolutional operation on firm features models spatial dependency among each firm, and the Gated Recurrent Units (GRU) (Chung et al., 2014) framework captures temporal patterns in each time series. An effective predictive model for cross-sectional return prediction requires learning both spatial and temporal dependency. The general network representation model mainly focuses on learning node embedding from networks and largely ignores the node features. Meanwhile, the importance of firm-specific factors in asset pricing is evident throughout finance literature. To resolve this, we use convolutional diffusion. The convolutional diffusion function allows us to incorporate firms' characteristics as node features into the learning process.

Our model uncovers several interesting findings by combining the attention and diffusion mechanism. First, the proposed model is superior in prediction accuracy to traditional asset pricing methods and other off-the-shelf machine learning models. Notably, the prediction errors in our model decrease by 6% compared to the state-of-the-art recurrent graph diffusion model (Li et al., 2018). Statistical tests like Diebold-Mariano (DM) and Giacomini-White (GW) show that the proposed model significantly outperforms other benchmark models. Second, the proposed model provides profitable investment strategies for investors. The Sharpe ratio of our model outperforms the S&P-500 index Sharpe ratio by 34% and the second-best

model N-BEATS (Oreshkin et al., 2019) Sharpe ratio by 24%. Like other benchmarks, our model is not immune to the market downside, but it still manages to select the stocks that recover faster than those chosen by all other benchmark models. With a monthly holding period, the long-short portfolio constructed based on DY-GAP prediction can earn an annualized return of 17%. Third, network information is essential for asset return predictability. Incorporating network information improves predictability by 24% compared to simple attention to features. Fourth, in addition to the network structure, the positive and negative signs of edge weights in a given network convey valuable properties of the network. In times of crisis, the positive linkage between firms rapidly increases; at the same time, negative edges disappear, and the network forms a ball-shaped structure, which is consistent with the findings of Sandhu et al. (2016). In addition, incorporating the signed information reduces predictive error by 20% over the unsigned model. The results are robust over multiple sensitivity tests, with different assessment metrics and synthetic datasets.

This paper contributes to the emerging literature that examines network learning and asset pricing in four ways. First, we propose an end-to-end asset pricing model that learns the firm network from data and captures the Spatio-temporal relationships for downstream prediction. Although Veličković et al. (2018) and Li et al. (2018) use graph learning and diffusion convolution separately in different settings, our approach combines both techniques to create a framework for tracking dynamic networks in the financial market. More importantly, our model is designed to extract dynamic graphs from time series without requiring pre-defined network topologies, whereas Graph Attention Networks (GAT) (Veličković et al., 2018) and Diffusion Convolutional Recurrent Neural Network (DCRNN) (Li et al., 2018) are implemented for (pre-existing) static graphs.

Second, we create separate learning modules for positive and negative connections and use neural networks to combine the outputs of both modules for return prediction. Most network techniques developed in other areas focus on unsigned graphs. However, the relationship among firms in the equity market can be positive or negative. A positive (negative) relation

between two firms signifies that the firms co-move in the same (opposite) direction in the marketplace. Figure 2 illustrates that signed networks contain valuable information for peer identification and clustering (Esmailian & Jalili, 2015). In addition, the presence of negative connections provides the basis for effective financial investment strategies, such as portfolio management and diversification. Sector similarity networks and customer-supplier networks cannot capture negative connections as they only contain adjacency edges with "1" for connected and "0" for not-connected. Correlation-based networks contain both positive and negative weights, but many off-the-shelf network techniques are not designed for this type of network. The common practice is still to use absolute values with the assumption that significant correlation represents high similarity regardless of the sign (Namaki et al., 2011; Tumminello et al., 2010; Di Cerbo & Taylor, 2020). This assumption violates the core idea that a positive correlation indicates convergence while a negative correlation indicates divergence. In contrast, our model learns both positive and negative connections separately at each time point and determines their contribution according to the edge significance.

Third, the proposed model is a novel application of graph neural networks in asset pricing and contributes to the emerging literature on machine learning-based asset pricing models. Unlike previous work, we use graph neural networks to integrate heterogeneous datasets, e.g., the fundamentals, historical returns, and the US equity market network structure, to improve asset price prediction. The experiments on both monthly and daily data reveal our model's superiority, robustness, and rationale of its predictive capacity.

Fourth, the proposed model is easily generalized and extendable in multiple contexts. In this paper, we applied the DY-GAP model to synthetic and real financial market data and documented significant performance improvements over multiple reference models. In addition, we demonstrate how easy it is to incorporate other characteristics or macroeconomic factors in the proposed model to account for changing external factors and improve the information content of the model.

The rest of the paper is organized as follows: Section 2 provides a brief review of related

literature on graph neural network, equity network representation, and machine learning applications in asset pricing. Section 3 provides the details of our proposed model. Section 4 describes the data and experimental settings. Section 5 discusses the empirical results. We perform an extensive robustness test in Section 6 that includes an ablation study, sensitivity study, and a generalized extension of the model with synthetic and macroeconomic data. Finally, Section 7 offers the conclusions and future research direction.

2. Related Literature

This paper is related to three streams of literature. The first one is the network study in computer science and information system. Several seminal works on graph and network representation learning initially appeared in computer science and are sequentially adopted by other domains. These efforts include spectral analysis in networks, node embedding, node classification, edge embedding, edge dynamics modeling, and their applications in various domains, including image processing, protein-protein interactions, and social network analysis (Post, 2012; Grover & Leskovec, 2016; Perozzi et al., 2014). The advancement in machine learning and deep learning paves the way for developing neural networks on large-scale complex graphs. Especially, a surge of spectral-based graph neural networks follows the seminal work of Kipf & Welling (2017). These Graph Convolutional Network models (GNN) focus primarily on static networks with predefined topologies and are mainly used for node classification and link prediction (Defferrard et al., 2016). A detailed review of the literature is presented in Wu et al. (2020). The success in network classification (predicting categorical values) motivates researchers to apply graph neural networks in regression problems (predicting continuous values). In recent years, several spatiotemporal models enhanced by graph neural networks attain impressive results in traffic prediction (Li et al., 2018; Fang et al., 2019; Diao et al., 2019), ride-hailing demand forecasting (Geng et al., 2019), and COVID-19 trend prediction (Cao et al., 2020). Motivated by these successes, we apply graph neural networks to forecast asset return in this paper.

The second stream is related to the network representation of firms. A handful of recent finance research activities attempt to understand the network dynamics of financial assets. These include developing graph representation for financial market (Tumminello et al., 2010; Herskovic, 2018; Di Cerbo & Taylor, 2020) and modeling information flow and shock transmission among financial assets (Elliott et al., 2014; So et al., 2020; Le et al., 2020) and the corresponding institutions (Rogers & Veraart, 2013). Among various techniques used for representing firm networks, the correlations of historical returns are the most dominant. This technique constructs the initial networks based on the Pearson correlation of historical returns and applies a threshold function or minimum spanning tree to sparsify the correlation network.

Because most popular graph techniques are designed for unsigned graph, to apply those techniques, researchers often use only positive correlations or absolute values of the correlations in these cases (Tumminello et al., 2010; Di Cerbo & Taylor, 2020). Sophisticated network models, such as dynamic conditional correlations and transfer entropy, are also used to identify information flow between financial markets (Dimpfl & Peter, 2013; Sensoy et al., 2017; Korbel et al., 2019) and instruments (Assaf et al., 2022). Cohen & Frazzini (2008) and Herskovic (2018) study the customer-supplier network at the industry level and construct new pricing factors from the network. Muslu et al. (2014) suggests that the stock price of one firm can be affected by other stocks with the same analysts' coverage. Grullon et al. (2014) argue that firms are affiliated with investment banks, and we can use the investment banks to produce an underwriter network. Studies also show that firms can be linked through managers who often have a rich set of social networks via prior employment, education, or memberships in social, cultural, and charity organizations (Hwang & Kim, 2009; Bruynseels & Cardinaels, 2014).

Those studies also provide essential insights on the significance of networks on the information flow between financial entities. A large body of literature documents the importance of network in contagion effects and systemic risks (Acemoglu et al., 2015; Glasserman &

Young, 2015; Elliott et al., 2014; Billio et al., 2016; Buraschi & Porchia, 2012; Rogers & Veraart, 2013; So et al., 2020). Carvalho (2014) shows that local shocks occurring in the production network might propagate across the economy and stimulate aggregated fluctuations. Ozdagli & Weber (2017) finds that the effects of monetary policy shocks are largely driven by the production networks, suggesting that the network is an important propagation mechanism of monetary policy to the real economy. Hou (2007) suggests that the structure of product markets affects managers' equilibrium operation decisions. These decisions affect the risk of a firm's cash flow and then stock prices.

The third stream is related to the studies on using machine learning models to predict asset prices. Predicting asset price is of great importance to academic research and real-world investment. Traditional asset pricing models mainly focus on uncovering risk factors (Fama & French, 2015; Kelly et al., 2019). Over the years, economists have identified hundreds of factors to explain the variability in returns among assets (Feng et al., 2020). The advancement in machine learning and deep learning allows researchers to use the full spectrum of these factors (for example, market, macroeconomic, and firm's accounting fundamentals) to predict returns or asset prices (Heaton et al., 2017; Gu et al., 2020a,b; Uddin & Yu, 2020; Zhang et al., 2017; Yu & Yan, 2020; Li et al., 2019). Among them, auto-encoder is applied in Gu et al. (2020a) and Uddin & Yu (2020) to forecast stock returns from historical data. Gu et al. (2020b) use Multi-layer-perceptron (MLP) on 94 characteristics variables to predict stock returns. Chen et al. (2020) use the generative adversarial network (GAN) in a non-linear asset pricing model based on the no-arbitrage condition.

Given the importance of networks in financial entities, incorporating network information in asset pricing is surprisingly limited. There are only a handful of studies that attempt to fill the gap and examine how the network enhances the predictability of future stock returns (Herskovic, 2018; Kim et al., 2019; Feng et al., 2019; Li et al., 2020). Our paper differs from these network-based approaches in several aspects. Feng et al. (2019) use temporal graph convolutional neural network to rank assets for portfolio optimization. Different from our

work, they use feature similarities to construct the initial graph. Herskovic (2018) proposes to use the multi-sector information to build a static input-output network. In contrast, our model captures the time-varying information and overcomes the drawbacks associated with the assumption of sector similarity. Li et al. (2020) propose the LSTM Relational Graph Convolutional Network to predict the overnight stock movements in a fixed firm correlation network that fails to capture the latent components of the network and the dynamic changes in the network. The stock prediction model in Kim et al. (2019) applies hierarchical attention in learning node representations from spatial convolutions without considering any temporal dynamics. On the other hand, our model benefits from the recurrent diffusion step and captures both spatial and temporal dependencies.

3. Methodology

In this section, we first define the problem of asset return prediction and then present the building blocks of the DYnamic Graph neural network for Asset Pricing (DY-GAP).

3.1. Problem Definition

The fundamental question in asset pricing is 'why do different assets earn different average returns?' Over the years, several models have been proposed to explain the return difference among assets. These models are grouped into two broader categories based on their inputs: (i) time-series models and (ii) factor models.

Time series models mainly use historical returns to forecast future returns. Researchers find that the historical average excess stock return is a better predictor of future excess stock returns (Welch & Goyal, 2008). The rationale of these findings is the time-varying pattern, momentum, and reversal effects in the stock return (Chordia & Shivakumar, 2002). Several time-series-based models, such as ARIMA, GARCH, and VAR, are applied to predict future returns (Pai & Lin, 2005; Dong et al., 2020; Ibrahim et al., 2020). These models use the

function of historical returns to estimate future returns:

$$\hat{y}_{t+1} = f_1(y_t, \cdots, y_{t-K})$$

where y is the observed excess returns and K is the window size of historical returns.[1]

Factor models assume asset returns can be expressed as a linear function of various macroeconomic, market, and security-specific factors. These factors include, but are not limited to volatility, trading volume, and accounting ratios (Fama & French, 1989; Campbell & Thompson, 2008; Kelly et al., 2019).

$$\hat{y}_{t+1} = f_2(x_1, \cdots, x_P)$$

where P is the total number of factors.

We formulate the asset pricing problem that combines these two types of inputs, the historical returns and accounting fundamentals, to capture the return variability. First, we define multivariate temporal graphs at time t as $\mathcal{G}_t = (V_t, A_t)$, where V_t is the set of firms (nodes) $|V_t| = N$ and $A_t \in \mathbb{R}^{N \times N}$ is a weighted adjacency matrix representing firms' quantitative proximity to their neighbors at time t. The value of A_{ijt} indicates the strength of the interdependence between nodes i and j at time t. The edge (interconnection) set E_t in the graph can be derived from A_t: $A_{ijt} = 0$ indicates that the firms i and j are independent to each other ($\epsilon_{ijt} \notin E_t$) at time t, while non-zero indicates an edge $\epsilon_{ijt} \in E_t$ between i and j at time t. The input signal on the graph at time t is $X_t = \{x_{ip}\} \in \mathbb{R}^{N \times P}$ and the prediction output on all N nodes is $\hat{Y}_{t+1} \in \mathbb{R}^{N \times 1}$, where the $\hat{\cdot}$ means a predicted value and N is the number of firms (nodes). Instead of using pre-defined networks, we use neural networks to learn the adjacency matrix A_t at time t from the historical returns $Y_t = \{y_{ik}\} \in \mathbb{R}^{N \times K}$, where K is the rolling window size of historical returns. The Y_t is a matrix of excess return

[1]Following standard academic finance literature (Gu et al., 2020a,b; Kelly et al., 2019) for return we use the conditional expected stock returns in excess of the risk-free rate: $y_{i,t} = r_{i,t} - rf_t$.

over risk free rate of the corresponding time step.

Given the observed returns of previous K timestamp $Y_t, \cdots, Y_{t-K}$ and graph signal X_t, the objective of Graph Neural Networks (GNNs) is to "connect the dots" and learn an effective network structure $\mathcal{G}_t$ at each time step and in the meanwhile to predict the next time-step $\hat{Y}_{t+1}$ with the integrated model of Graph Neural Networks and Recurrent Neural Networks. We use neural network to implement the forecasting function $f(\cdot)$ with parameters $\Theta = \theta_1, \theta_2, \theta_3$ (θ_1 represents the neuron weights for "connecting the dots", θ_2 contains the parameters for the graph filter g_{θ_2} defined on $\mathcal{G}_t$, and θ_3 represents the neuron weights for transforming the input node features.):

$$\begin{aligned} \mathcal{G}_t &= Attn(Y_t, \cdots, Y_{t-K}; \theta_1) \\ \hat{Y}_{t+1} &= f(g_{\theta_2} \star X_t; \theta_2, \theta_3) \end{aligned} \tag{1}$$

3.2. DY-GAP Model Framework

Figure 3 presents the architecture of our proposed model. The model consists of three different learning functions (i) Embedding learning, (ii) Graph learning, and (iii) Spectral and temporal dynamics learning. We follow Bahdanau et al. (2015) and apply the self-attention function to learn the dynamic network structures from the historical return data. The return data is often noisy and large. An embedding learning layer is used to clean the noisy data and obtain the firm's condensed representation before performing the attention function. Finally, the spectral-temporal recurrent convolution function using the learned networks is performed on the firm fundamentals.

3.2.1. Embedding Learning

Principal Component Analysis (PCA) is widely used in finance for dimension reduction and feature extraction from time-series data (Bikbov & Chernov, 2010). At each time t, we perform Singular Value decomposition (SVD) on historical return and extract principal components for each firm. For simplicity, we drop the time index for the remaining discussion.

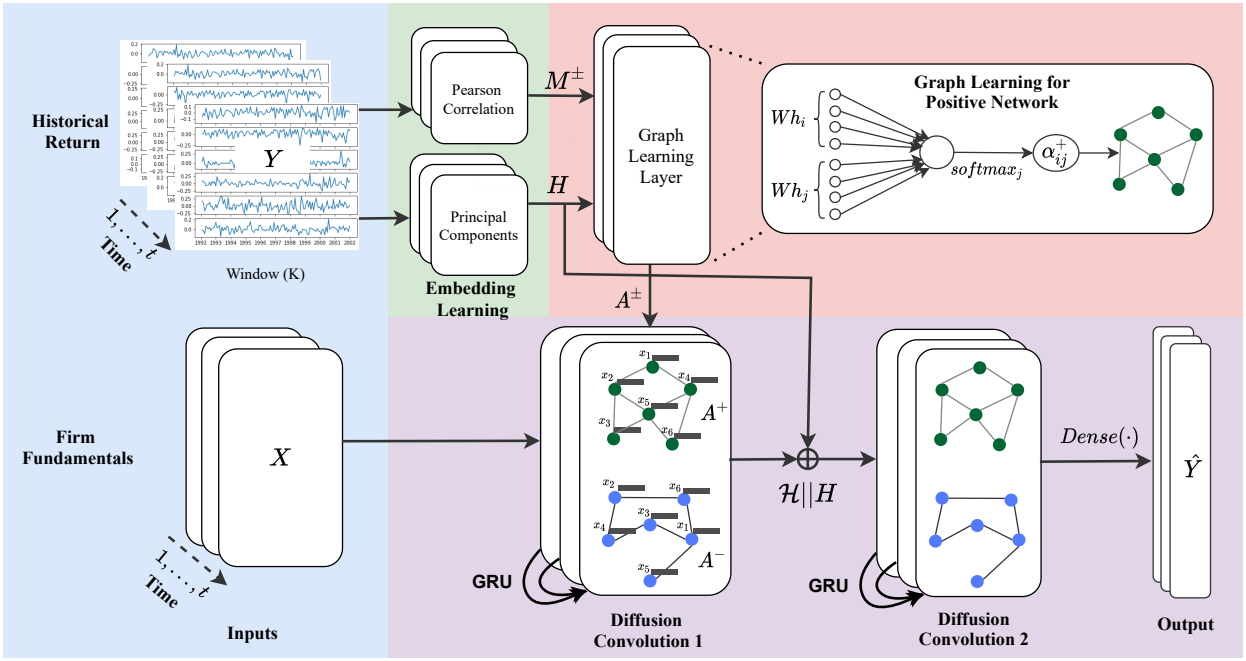


Figure 3: The three-stage DY-GAP model architecture. (1) SVD on the historical return learns latent embedding for each firm. (2) Self-attention on the latent embedding learns the network architecture. Two different attention blocks (The attention head for the positive network is shown in the upper right corner of the image) are performed to learn both positive (green) and negative (blue) networks. The Pearson correlation of historical return ensures efficient masked attention. (3) A diffusion convolution with the firm signals is applied to the learned network to discover the spatial dependency and the GRU recurrent neural network for temporal patterns. Two diffusion modules are used here, with the first on the firm fundamentals and the second on the concatenated embedding of the output of the first diffusion layer and the latent embedding from SVD. Solid arrows indicate the flow of information.

Given $Y \in \mathbb{R}^{N \times K}$, we perform SVD on Y as follows:

$$Y = USV^{\top} \tag{2}$$

where U is a unitary matrix, S is the diagonal matrix of singular values corresponding to the Eigenvalues in the correlation matrix, and $H = US$ consists of the principal components. Assume that the input matrix has a low rank, we perform dimension reduction and choose the first L principal components $\sum_{i=1}^{L} U_{:i} S_{:i}^{\top}$ as the embedding matrix of return data on each time t, where $L < K$. With N firms, $H_t = \{h_{ilt}\} \in \mathbb{R}^{N \times L}$ holds the initial node embeddings (features) for firms at time t.

3.2.2. Graph Learning

The graph learning layer employs the attention mechanism proposed in Bahdanau et al. (2015) to calculate the attention coefficients between each target node and its neighbors and learn the edge weights between two neighboring nodes. At each time point, the attention function is performed on each firm's embedding H_t learned by the previous SVD function. Every node (firm) attends to its neighboring nodes embedding, given its own embedding as the query. We use neural network to calculate attention score of each node in relation to others. Neural network facilitates learnable parameters (learnable vector of parameters $\vec{a}$ and learnable linear projection matrix W) to expand modeling capacity. In particular, we follow the Graph Attention Networks (GAT) of Veličković et al. (2018): first applying a linear transformation with weight matrix $W \in \mathbb{R}^{L' \times L}$ on each node of the attention pair and then attending the transformation results with the shared attention weight vector $\vec{a}$: $\mathbb{R}^{L'} \times \mathbb{R}^{L'} \to \mathbb{R}$. The learned attention coefficient is:

$$e_{ij} = \vec{a}(\mathbf{W}h_{i:}, \mathbf{W}h_{j:}). \tag{3}$$

The learned value e_{ij} indicates the importance of firm j's return on that of firm i. We adopt an activation function to add nonlinearity to the dependence relations among firms and apply the softmax function to normalize coefficients and make them easily comparable across nodes. The final network edge weight α_{ij} is defined as follows:

$$\alpha_{ij} = \frac{\exp(\sigma(\mathrm{a}^\top[\mathbf{W}h_{i:}||\mathbf{W}h_{j:}]))}{\sum_{j=1}^{n}\exp(\sigma(\mathrm{a}^\top[\mathbf{W}h_{i:}||\mathbf{W}h_{j:}]))} \tag{4}$$

where $\top$, $||$, and σ represent transpose, concatenation and nonlinear activation operation, respectively. $\mathrm{a} \in \mathbb{R}^{2L'}$ is a trainable weight vector parametrizing the attention mechanism $a(\mathbf{W}h_{i:}, \mathbf{W}h_{j:})$. The softmax operation ensures $\sum_{j=1}^{n}\alpha_{ij} = 1$ and $\alpha_{ij} >= 0$ and thereby, normalizes node degree across the entire learned network.

Veličković et al. (2018) suggest that the desired objective can be achieved by a simple

attention strategy of attending to all node pairs indiscriminately while ignoring any pre-existing graph structure. However, as discussed in Section 1, the interactions among firms are not uniform and rather complex: some of them converge and render the market in the same direction, while the others diverge and lead to heterogeneous behavior in the market. Later we show in Section 5.4 that the distribution of the interaction relationships aligns with the market conditions.

A firm does not necessarily connect to all other firms, and a spurious relationship might do more harm than no connection due to the overfitting problem and excessive computation costs. Therefore, to expedite the graph learning process, we use the masked attention function that only considers the closely relevant firms of a target firm. We use the Pearson correlation coefficients of the historical return data to determine a firm's initial connections and its potential neighbors.[2] The Pearson's correlation, ρ_{ij}, between firm i and j at time t with a rolling window K is defined as follows:

$$\rho_{ij} = \frac{Cov(r_i, r_j)}{\sqrt{Var(r_i)Var(r_j)}}.$$

We keep the rolling window size the same as the one in the embedding learning module for both daily and monthly data. The original return correlation matrix is dense, with almost all firms' returns (positively or negatively) being correlated to other firms' returns. The values close to zero do not provide meaningful information about firms' similarities but noise. To enhance the signal-to-noise ratio (SNR), we adopt the noise filtering technique proposed in Di Cerbo & Taylor (2020). The signal enhancement algorithm first denoises the empirical Pearson covariance matrix by performing an eigendecomposition and replacing the

[2]We acknowledge that there are many techniques available for identifying initial connections among firms, and some are more sophisticated models than Pearson correlation, e.g., Dynamic Conditional Correlations (Sensoy et al., 2017) and Transfer Entropy (Korbel et al., 2019). We use a simpler Pearson correlation to demonstrate the importance of the network in the stock return prediction. The initial network from Pearson correlation can be easily replaced by a more sophisticated approach, such as, Granger causality and the transfer entropy to masked the attention function in the graph learning layer.

noisy eigenvalues with their average to preserve the trace of the correlation matrix. The denoised matrix might not be a positive definite symmetrical matrix. Finally, we apply convex optimization to enhance the signal and construct a correlation matrix that is the closest to the denoised matrix.

The original and denoised correlation matrices have both positive and negative coefficients. A positive correlation between i and j indicates that i and j reassemble to each other, while a negative one means the opposite. A single attention head might treat both positive and negative coefficients indiscriminately and violates the correlation property. Therefore, to learn meaningful network representations, we introduce two separate attention heads, each attending the positive and negative components of the firm correlation matrix separately. We first decompose the new correlation matrix into two mask matrices: M^+, where $M^+_{i,j} = 1$ if $\rho_{i,j} \geq 0$, otherwise 0; and M^- where $M^-_{i,j} = 1$ if $\rho_{i,j} < 0$, otherwise 0. Then we perform the following masked attentions according to the positive and negative matrices:

$$\alpha^{\pm}_{ij} = \frac{M^{\pm}_{ij}\exp(\sigma(\mathrm{a}^\top[\mathbf{W}h_{i:}||\mathbf{W}h_{j:}]))}{\sum_{j=1}^{n} M^{\pm}_{ij}\exp(\sigma(\mathrm{a}^\top[\mathbf{W}h_{i:}||\mathbf{W}h_{j:}]))}. \tag{5}$$

The learned attention coefficients $\alpha^{\pm}_{ij}$ are assembled into two new affinity matrices $A^+ = \{\alpha^+_{ij}\} \in \mathbb{R}^{N\times N}$ and $A^- = \{\alpha^-_{ij}\} \in \mathbb{R}^{N\times N}$. During the graph learning phase, the model learns $2T$ networks (T of A^+ and T of A^-) and each $A^{\pm}$ associated with a specific time point t.

3.2.3. Spectral and Temporal Dynamics Learning

We use the recurrent diffusion function to model the spatial relationship (network) and temporal dependency among firm fundamentals based on the learned network structure. We follow the graph convolutional diffusion process proposed in Li et al. (2018) to propagate firms' information to their neighbors.

Graph Diffusion. We apply two separate diffusions in two networks $A^{\pm}$ learned from the attention step. Because all node degrees equal to one, the diffusion process consists of a random walk on Graph $\mathcal{G}$ with the state transition matrix (A^+ or A^-) and its random walk

normalization $D^{\pm^{-1}}$, where $D^{\pm}$ is the diagonal matrix of the node degree, $D^{\pm}_{ij} = deg^{\pm}(v_i)$ if $i = j$, otherwise 0.[3] Following Li et al. (2018), we define the diffusion convolution as a set of the graph filters f_θ applied to each input channel $X_{:,p} \in \mathbb{R}^N$ $(1 \leq p \leq P)$ in the multivariate graph signal of firm fundamentals as follows:

$$f_\theta \star_{\mathcal{G}} X_{:,p} = \sum_{s=0}^{S-1} \theta_{s,1} (D^{+^{-1}} A^+)^s X_{:,p} \| \sum_{s=0}^{S-1} \theta_{s,2} (D^{-^{-1}} A^-)^s X_{:,p} \tag{6}$$

where $\theta \in \mathbb{R}^{S \times 2}$ are the parameters for filter and $D^{+^{-1}} A^+$, $D^{-^{-1}} A^-$ are the (random-walk normalized) positive and negative transition matrices of the diffusion process, respectively. S denotes the number of diffusion steps. Li et al. (2018) show that a sufficiently large number of diffusion will converge to a stationary distribution. It is important to note that, unlike the work in Li et al. (2018) that deals with directed graphs, our model consists of two undirected graphs with positive and negative edges, performs diffusion on these two networks that correspond to the first and second term in Eqn 6, and concatenates the two outputs. The two-term diffusion allows us to learn the spatial dependencies from the positively and negatively connected neighborhood. The diffusion convolution layer consists of Q graph filters and takes $X \in \mathbb{R}^{N \times P}$ as input and output $\mathcal{H} \in \mathbb{R}^{N \times 2Q}$ where P is the number of input channels and Q is the number of output channels for each diffusion network. The concatenation function of positive and negative diffusions in Equation 6 results in doubling the output features Q. For each output channel $q \in 1, \ldots, Q$, the output is as follows:

$$\mathcal{H}_{:,q} = \sigma \left(\sum_{p=1}^{P} f_{\Theta_{p,q,:,:}} \star_{\mathcal{G}} X_{:p} \right) \tag{7}$$

where $\Theta_{p,q,:,:} \in \mathbb{R}^{S \times 2}$. $\{f_{\Theta_{p,q,:,:}}\}$ are the diffusion filters and σ is the activation function, i.e., ReLU.

[3] In our case, all row sums in $A^{\pm}$ equal to 1, and degree matrices $D^{\pm}$ become an identity matrix. We still keep $D^{\pm}$ in the subsequent convolutional diffusion equations and make them generalize to any other un-normalized affinity matrix $A^{\pm}$.

GRU for Temporal Modeling. In the asset pricing model, the volatility of past prices bears significant insight into the future probability distribution of prices. Furthermore, it is essential to model the temporal dependency along with the spatial dependence. Gated Recurrent Units (GRU) (Chung et al., 2014) is proved to be effective for temporal dependency modeling. As suggested in Li et al. (2018), we replace the common linear transformations to the inputs and hidden states for all GRU gates with the graph diffusion operation defined in Equation 7 and augment the GRU with the diffusion convolution. The updated GRU function $\mathcal{H}_t = GRU_\Theta(X_t, \mathcal{H}_{t-1})$ is defined as follows:

$$\begin{aligned} r_t &= \sigma(\Theta_r \star_\mathcal{G} [X_t, \mathcal{H}_{t-1}] + b_r) \\ u_t &= \sigma(\Theta_u \star_\mathcal{G} [X_t, \mathcal{H}_{t-1}] + b_u) \\ C_t &= \sigma(\Theta_c \star_\mathcal{G} [X_t, (r^t \odot \mathcal{H}_{t-1})] + b_u) \\ \mathcal{H}_t &= u_t \odot \mathcal{H}_{t-1} + (1 - u_t) \odot C_t \end{aligned} \tag{8}$$

where $X_t, r_t, u_t, \mathcal{H}_t$, denote the input, reset, update gate and the hidden state at time t, respectively. $\star_\mathcal{G}$ denotes the graph convolution diffusion and $\Theta_r, \Theta_u, \Theta_c$ contain the filter parameters of the reset gate, update gate, and candidate activation respectively.

The embedding H_t serves two roles in the proposed framework in Figure 3: learning attention coefficients and supplying the stock market information of each firm. To support the second role, we design two recurrent diffusion convolution layers to process the firm signals in Figure 3. The first layer performs the diffusion convolution on the firm fundamentals and the hidden state $[X_t, \mathcal{H}_{t-1}]$. The second layer concatenates the output $\mathcal{H}_t$ from the first diffusion layer and the embedding H_t from the embedding learning module and applies another recurrent diffusion on $\mathcal{H}_t || H_t$. This concatenation function implements a bypass and ensures that the return data is directly incorporated into the convolution framework. Consequently, the bypass design avoids the vanishing gradient problem and significantly improves the learning process. The recurrent diffusion using GRU defined in Equation 8 is also applied in the

second layer, where the diffusion convolution is expressed as follows:

$$\mathcal{H}'_{:,q,t} = \sigma\left(\sum_{p=1}^{P} f_{\Theta_{p,q,:,:}} \star_{\mathcal{G}} [\mathcal{H}_{:,p,t}||H_{:,p,t}]\right) \tag{9}$$

Finally, to predict returns, the diffusion network uses a dense layer that regresses the prediction output $\hat{Y}_{t+1} \in \mathbb{R}^{N \times P}$ onto the hidden state of the second recurrent diffusion layer. A dense layer is a linear layer without any activation function.

$$\hat{Y}_{t+1} = f(\mathcal{H}'_t; \theta') \tag{10}$$

where $\mathcal{H}'_t$ is the output in the final GRU layer of the second diffusion, and θ' contains the learnable weights of the dense layer of multilayer perceptron (MLP).

4. Data and Model Setting

4.1. Data

We use two sets of stock data with different frequencies – monthly and daily. The monthly data includes stocks in the Russell 3000 index from January 01, 1990 to December 31, 2019. It is divided into three sets: training (January 1990 to December 2009), validation (January 2010 to December 2013), and test (January 2014 to December 2019). The daily data are stocks in the S&P-500 index. The sample period is from January 01, 2010, to December 30, 2020. It is divided into training, validation, and test sub-samples as well, which are January 01-2010 to December 31-2015, January 01-2016 to December 31-2017, and January 01-2018 to December 31-2020, respectively.[4] To be included, a given firm must have at least two-years of financial information and must exist during the full sample periods. As monthly data has longer periods, more firms are removed and the final sample includes 1098 stocks from the Russell 3000 index in the monthly data and 452 stocks from the S&P-500 index in the daily

[4]For both monthly and daily data, firms are selected based on the index constituents by October 31, 2020.

Table 1: Fundamentals Variables

Firm Characteristics	Calculation Procedure
Total assets to market	Total asset / market value of equity
Size	log $(p_t \times share\ outstanding)$
Turnover	Volume / share outstanding
Growth rate of volume	$V_t - V_{t-1}/V_{t-1}$
Growth rate of share outstanding	$SO_t - SO_{t-1}/SO_{t-1}$
Closeness to past year high	$P_{t-1} - max(P_{t-1}, ...p_{t-12})/max(p_{t-1}, ..., p_{t-12})$
Closeness to past year low	$P_{t-1} - min(P_{t-1}, ...p_{t-12})/min(p_{t-1}, ..., p_{t-12})$
Spread	$p^h_{t-1} - p^l_{t-1}$, monthly [daily] high minus low price
Opening and closing spread	$(p^o_{t-1} - p^c_{t-1})$, daily opening minus closing price
Capital gain	Value is 0, if no capital gain is recorded
EPS	Earning per share
Dividend	Dividend paid in cash
Total volatility	Price volatility of last 60 months [last 252 days].
Idiosyncratic volatility	Total volatility - market volatility
Market return	Return on S&P-500 index
CAPM market Beta	Beta on Fama-French market factor
Small minus big beta	Beta on Fama-French size factor
High minus low beta	Beta on Fama-French value factor
1 week momentum	$(p_{t-1} - p_{t-5})/p_{t-5}$ -daily only
2-week momentum	$(p_{t-1} - p_{t-10})/p_{t-10}$ -daily only
1-month momentum	$(p_{t-1} - p_{t-2})/p_{t-2}$ $[(p_{t-1} - p_{t-21})/p_{t-21}]$
2-month momentum	$(p_{t-1} - p_{t-42})/p_{t-42}$ -daily only
3 month momentum	$(p_{t-1} - p_{t-3})/p_{t-3}$ $[(p_{t-1} - p_{t-63})/p_{t-63}]$
6 month momentum	$(p_{t-1} - p_{t-6})/p_{t-6}$ $[(p_{t-1} - p_{t-126})/p_{t-126}]$
12-month momentum	$(p_{t-1} - p_{t-12})/p_{t-12}$ -monthly only

data. The monthly and daily stock returns are from CRSP, and firms' fundamental variables are from Compustat. Due to the data availability, the fundamental variables included in the daily and monthly analysis are slightly different. Following the asset pricing literature, we incorporate 21 monthly variables and 24 daily variables (Fama & French, 1989; Chordia & Shivakumar, 2002; Campbell & Thompson, 2008). Table 1 provides the details of these variables.

Table 2: Hyperparameter Settings

Hyperparameter	Values
Monthly window size (months)	12, 24, 36, 48
Daily window size (trading days)	44, 66, 122, 252
Principal components	5, 10, 15
Output channels	8, 16, 32
Attention heads	6, 8, 10, 12
Diffusion steps	6, 8, 10, 12

4.2. DY-GAP Setting

The DY-GAP model is implemented in Python using Keras (Chollet et al., 2015) with the TensorFlow in the back-end.[5] Keras is a popular open-source deep-learning API that allows more straightforward implementation of complex machine learning models than the classic Google Tensorflow. It also supports mini-batch gradient descent, multi-GPU, and distributed training. As a result, the training process converges faster.

The Hyperparameters of the experimental setup are reported in Table 2. The initial values of these hyperparameters are selected based on the model requirements and commonly used values in the literature. We choose the best hyperparameters based on the performance of the validation data. On the final selected hyperparameters, the hidden dimensionality in the principal component analysis-based embedding layer is $L = 10$. For monthly data, the window size K of historical returns for the embedding layer and Pearson correlation is 36 months; for daily data, it is 122 trading days (6 months). In the attention layer, the final selected model have 16 output channels with eight attention heads, and for diffusion convolution layers, 16 diffusion output channels with ten diffusion steps. To avoid overfitting, early stopping criteria are adopted for training the DY-GAP model. According to these criteria, the model stops training once the validation loss does not decrease for ten epochs (Li et al., 2018; Oreshkin et al., 2019; Veličković et al., 2018).

[5]The source code for DY-GAP with data will be shared in the published version for reproducibility.

4.3. Baselines

To assess the prediction performance of model, we first compare our model with multiple baseline models including multi-factor pricing models and time-series methods. For multi-factor pricing models, we consider the well-acknowledged Fama-French five-factor model (Fama & French, 2015), multivariate regression with fundamentals (Adrian et al., 2015), and recently proposed Empirical Asset Pricing via Machine Learning (EAP-ML) (Gu et al., 2020b) model. For time-series models, we use the classic ARIMA and several advanced deep neural network-based approaches, such as fully connected Long short-term memory (FC-LSTM) (Sutskever et al., 2014) and the state-of-the-art Neural Basis Expansion Analysis for interpretable Time series forecasting (N-BEATS) (Oreshkin et al., 2019). We also consider classic machine learning based models: Support Vector Regression (SVR) (Drucker et al., 1996) and Random Forest (RF) (Ho, 1995). Lastly, as our approach is inspired by two models, Graph Attention Network (GAT) (Veličković et al., 2018) and diffusion convolution recurrent neural network (DCRNN) (Li et al., 2018). We compare the performance of our model with them and show the advantage of integration.

We implement all machine learning-based models using Keras or Pytorch. For most models, we adopt the authors' source codes, if available, with necessary modifications. The best hyper-parameters are chosen based on the validation dataset. For MR, FF-5, and ARIMA, we learn the coefficients with the training and validation data, and then use the learned coefficients to estimate the performance for test data. Detailed descriptions of baseline models and their parameters are presented in Appendix A.

4.4. Evaluation Metrics

Three performance metrics, RMSE (Root Mean Square Error), MAE (Mean Absolute Error), and MAPE (Mean Absolute Percentage Error), are applied to evaluate the return prediction performance of models. The construction of those three metrics are as follows:

$$RMSE = \sqrt{\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}(y_{it} - \hat{y}_{it})^2} \tag{11}$$

$$MAE = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}|y_{it} - \hat{y}_{it}| \tag{12}$$

$$MAPE = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\left|\frac{(y_{it} - \hat{y}_{it})}{y_{it}}\right| \times 100 \tag{13}$$

where, $\hat{y}_{it}$ is the predicted excess return for firm i at time t and y_{it} is the actual excess return over risk free rate of the firm at that time. The small values of these metrics represent low prediction errors and high accuracy.

5. Empirical Results

In this section, the performance of our model is evaluated from three perspectives. First, we calculate the prediction errors of each model and show that our model consistently has the smallest forecast errors. Second, portfolios are constructed based on the predicted returns. If the prediction of a given model is more accurate, the portfolio based on the model more likely generates positive and higher returns. Third, we examine the network learning capacity of our model, especially around major financial and economic events.

5.1. Return Prediction

The model performance is first evaluated in terms of the prediction accuracy of future returns. Table 3 shows the prediction errors (RMSE, MAE, and MAPE) of each model, where panel A is for monthly data (Russel-3000 Index) and panel B for daily data (S&P-500 Index). Results are based on the test data sets only. The stochastic nature of machine learning models may lead to different forecasts with different initializations. To ensure the stability of our proposed model, we run each model ten times with different random seeds and report the average value with one standard deviation.

Table 3: Prediction Results and Significance Tests

	RMSE	MAE	MAPE (%)	DM-stat	GM-stat
Panel A: Monthly Data					
MR	0.1211	0.0944	66.4063	31.6490***	306.2051***
FF-5	0.0941	0.0843	58.9432	11.2842***	150.8288***
ARIMA	0.1141	0.0852	62.8968	29.9470***	295.7521***
EAP-MLP	0.0969±.009	0.0726±.007	61.5435±7.3	12.0029***	136.3969***
SVR	0.0941±.004	0.0751±.003	59.8505±2.1	7.0632***	45.9686***
RF	0.0921±.002	0.0739±.001	55.9597±1.0	7.9013***	10.0131***
FC-LSTM	0.0949±.006	0.0723±.006	56.1883±4.2	11.9735***	56.3103***
N-BEATS	0.1065±.001	0.0739±.001	61.4771±2.3	25.2504***	144.7827***
DCRNN	0.0912±.002	0.0727±.001	54.8688±1.2	1.7827**	6.3146**
GAT	0.0953±.006	0.0778±.005	61.7650±5.5	11.7096***	8.7344**
DY-GAP	0.0853±.002	0.0632±.001	52.2400±1.3		
Panel B: Daily Data					
MR	0.0310	0.0251	24.5875	11.3993***	110.8296***
FF-5	0.0274	0.0209	20.8055	10.7619***	101.0878***
ARIMA	0.0351	0.0252	32.6051	11.1121***	107.2008***
EAP-MLP	0.0519±.018	0.0524±.017	42.0046±9.1	66.5957***	642.8999***
SVR	0.0330±.003	0.0247±.004	24.3174±2.1	26.1884***	355.9468***
RF	0.0296±.007	0.0189±.006	17.2485±1.9	7.3564***	49.4854***
FC-LSTM	0.0254±.003	0.0183±.001	19.6615±0.9	11.3390***	110.8354***
N-BEATS	0.0241±.008	0.0173±.002	17.3975±1.1	4.3249***	11.8571***
DCRNN	0.0230±.005	0.0178±.002	16.8916±1.0	1.3157	1.1493
GAT	0.0292±.006	0.0195±.005	17.4611±1.6	1.8740**	8.5135***
DY-GAP	0.0233±.003	0.0161±.001	15.6098±0.7		

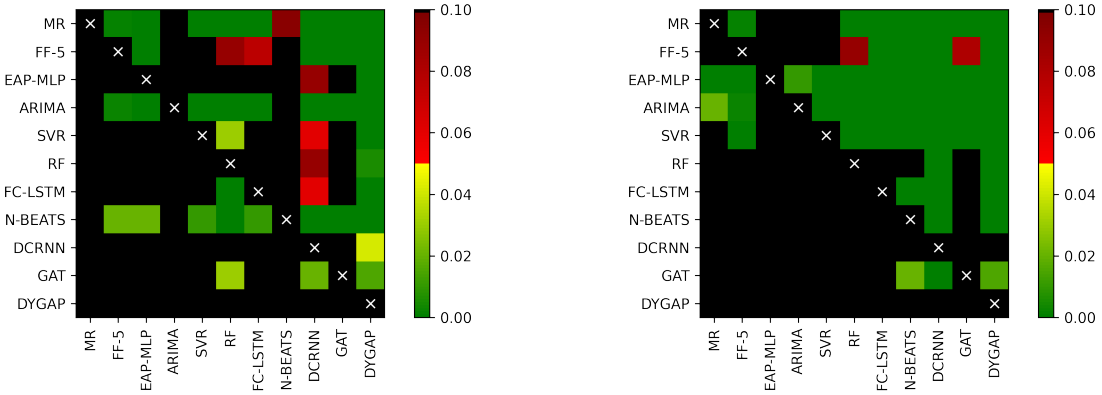
Note: This Table reports the return prediction results from our proposed model (DY-GAP) and other benchmarks. Panel A reports results from the monthly data (Russel-3000), and panel B report results from the daily data (S&P-500). Reported results are from test data sets only. The first three columns report the evaluation metrics RMSE, MAE, and MAPE. A lower value of RMSE, MAE, and MAPE indicates better performance. Columns 4 and 5 report the t-statistics of the Diebold-Mariano test and Giacomini-White (GW) test with the null hypothesis that the mean absolute forecast error of the "benchmark model" is smaller than or equal to that of "DY-GAP". Rejecting the null hypothesis indicates that the prediction of DY-GAP is significantly more accurate than a given benchmark. ***, **, * indicate to reject the null hypothesis at the 1%, 5%, and 10% significance level, respectively.

Our proposed model DY-GAP outperforms all other models in each performance matric using monthly data. In terms of RMSE, DY-GAP beats FF-5 and EAP-MLP by 9% and 12%, respectively. The two best performing models among all machine learning-based baseline models are FC-LSTM and DCRNN. Nevertheless, DY-GAP still outperforms these two models by 10% and 6% in the RMSE, and 7% and 5% in MAPE, respectively.

When using daily data, the superior performance of DY-GAP in prediction remains only with one exception, where DCRNN is the best model in terms of RMSE. However, the difference between DCRNN and our DY-GAP is only 0.0003 and statistically insignificant, and the performance of DY-GAP is more stable with a smaller standard deviation. In the next section, we further show that DCRNN is not stable with high volatility. The prediction error of DY-GAP in MAPE is smaller than DCRNN by 7%, GAT by 11%, N-BEAT by 10%, and FF-5 by 25%. In addition, the superiority of sophisticated predictive models becomes visible in the daily data. As the data frequency and size increase, sophisticated models take advantage of large datasets to minimize their overfitting problem and use their increased learning capacity to reduce bias. As a result, the state-of-the-art N-BEATS outperforms the earlier deep learning model the FC-LSTM by 5% (RMSE) in the prediction with daily data and becomes the third-best model.

We also examine whether the difference in forecast errors is significant or due to the specific choice of data values in the sample, by performing Diebold-Mariano Test (DM) (Diebold & Mariano, 2002) and Giacomini-White test (Giacomini & White, 2006). Results are reported in columns 4 and 5 in Table 3. DM and GW are two most commonly used tests to determine whether the two forecasts are significantly different. While the DM test measures the unconditional predictive ability, the GW tests for the conditional predictive ability between two forecasts.[6] Here, we apply the one-sided version of the DM and GW test. The null hypothesis of DM test is that the mean of the loss differential of benchmark model is lower than or equal to that of DY-GAP. And the null hypothesis of GW test is that the

[6]For more details please see Diebold & Mariano (2002) and Giacomini & White (2006).



(a) Monthly Data (b) Daily Data

Figure 4: p-values from Giacomini-White Test. The null hypothesis of the GW-test is that the conditional predictive ability of the loss differential of models on the vertical axis is higher than or equal to that of the models on the horizontal axis. Rejecting the null hypothesis indicates that the predictive ability of the model in horizontal axis is significantly more accurate than the model on vertical axis. The color spectrum represents the p-value of the t-stat. Black (p-value > 0.10) indicating that the null hypothesis is not rejected at 10% significant level, red ($0.05 <=$ p-value < 0.10) represent the null hypothesis is rejected at 10% significant level, the yellow to green (p-value < 0.05) represent the null hypotheses is rejected at 5% and lower significant level.

conditional predictive ability of the loss differential of the benchmark models being higher or equal than that of the DY-GAP. Rejecting the null hypothesis in both these tests means that the prediction performance of DY-GAP is significantly better than those of benchmark models, e.g., MR, ARIMA, and SVR.

Figure 4 presents the heat-map of the p-values from the GW-test for (a) monthly and (b) daily data and provides an overall snapshot of all models and their predictive accuracy. Rejecting the null hypothesis of GW-test indicates that the predictive ability of models in the horizontal axis is significantly more accurate than models on the vertical axis. For the monthly data, RF, DCRNN, and DY-GAP show better accuracy in return prediction compared to other models. Among all the competing models, in total RF outperform seven, DCRNN outperform nine, and DY-GAP outperform all ten compared models. For the daily data, the three best models are N-BEATS, DCRNN, and DY-GAP. Among all the competing models, in total N-BEATS outperform seven, DCRNN outperform nine, and DY-GAP outperform nine competting models. Overall, these results signify that the prediction

accuracy of DY-GAP is significantly more accurate than other existing asset pricing and machine learning-based models.

The integration of two powerful methods and nonlinearity explain the superior performance of DY-GAP. The combination of attention and diffusion function allows us to harness the advantage of both GAT and DCRNN and therefore outperforms the independent applications of individual methods. Standalone GAT learns the network from the historical return and firm fundamentals and attempts to model the returns from the perspective of the network. Nevertheless, it does not use any recurrent function to model temporal dependency. On the other hand, DCRNN uses a prior network topology and models spatial and temporal dependencies based on this single static network. In contrast, our model DY-GAP learns dynamic network topology using attention function and then uses diffusion convolution to model spatial and temporal dependency. DY-GAP also has clear superiority over time series models, such as ARIMA, FC-LSTM, and N-BEATS. Because DY-GAP considers both time series and firm characteristics. On top of that, these time-series models are applied to individual firms. Information from the market or related firms is neglected in these models. Compared with multi-factor pricing models, such as FF-5 and EAP-MLP, our model considers the spatial connectedness and nonlinear interactions among asset returns and attains a significant performance margin.

5.2. Portfolio Performance

In this section, we further evaluate the prediction performance of each model and show the outperformance of our DY-GAP with portfolio analysis. If the prediction of a given model is more accurate, the portfolio based on the model likely earns positive and higher returns. For the monthly data, we take a long position on the top 10% stocks with the highest predicted returns at the beginning of each month, hold the position until the end of the month, and then liquidate (i.e., sell all stocks). Figure 5 reports the cumulative returns on all portfolios from January 2014 to December 2019. The thick blue line represents the cumulative return of our model in the test period, and the thick gray line is for the S&P-500

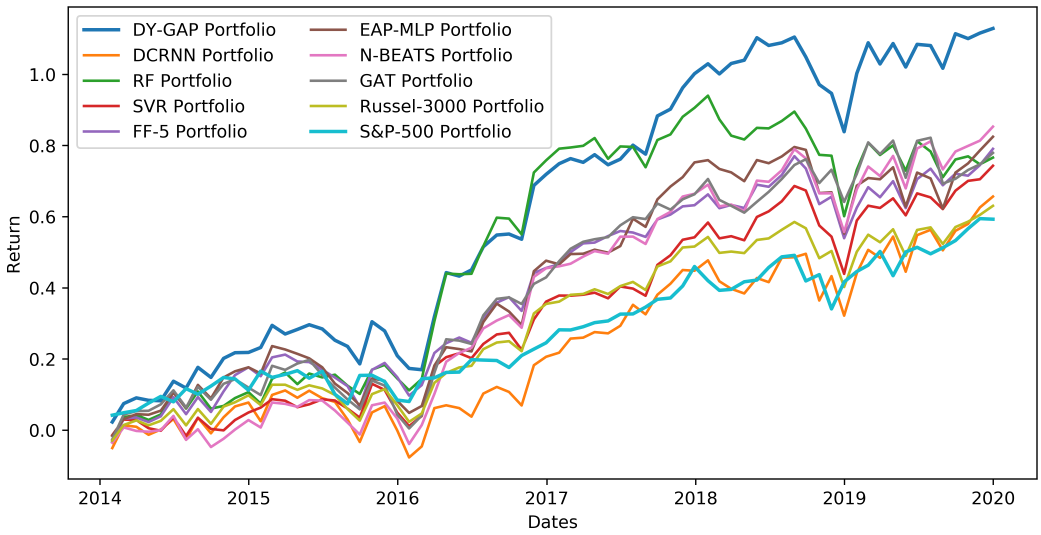


Figure 5: The portfolio performance from the monthly data. At the beginning of each month, we hold the top 10% stocks based on the prediction of each model and at the end of the month, we liquidate all stocks. The thick blue line is the cumulative return of our DY-GAP model in the test period and thick gray line is the S&P-500 index return during the test period. Russel-3000 is the average return of all our studied firms.

index return during the test period. Although the performance of the DY-GAP portfolio is indistinguishable from others during the early period, our model's superiority gradually stands out as time evolves. The cumulative return of our model doubles the S&P-500 index return and the weighted average return of all stocks in the Russel-3000 index. The finding signifies a profitable investment strategy that applies our model to identify the "success" group (i.e., the top 10% stocks with the highest returns). Among all baseline models, EAP-MLP and N-Beats also perform well in the cumulative returns.

The left panel in Table 4 reports the monthly average rate of return, monthly standard deviation, and annualized Sharpe ratio of different portfolios. Our model earns a high return rate (1.5% monthly) with relatively low risk (3%). Although DCRNN shows a high predictive power in Section 5.1, the portfolio analysis shows that the risk associated with DCRNN is much higher. As a result, the Sharpe ratio of DCRNN is lower than our model. S&P-500 index portfolio has the lowest risk, but with a low rate of return. The Sharpe ratio of the S&P-500 index portfolio is similar to that of MR, FF-5, ARIMA, and LSTM portfolios. The

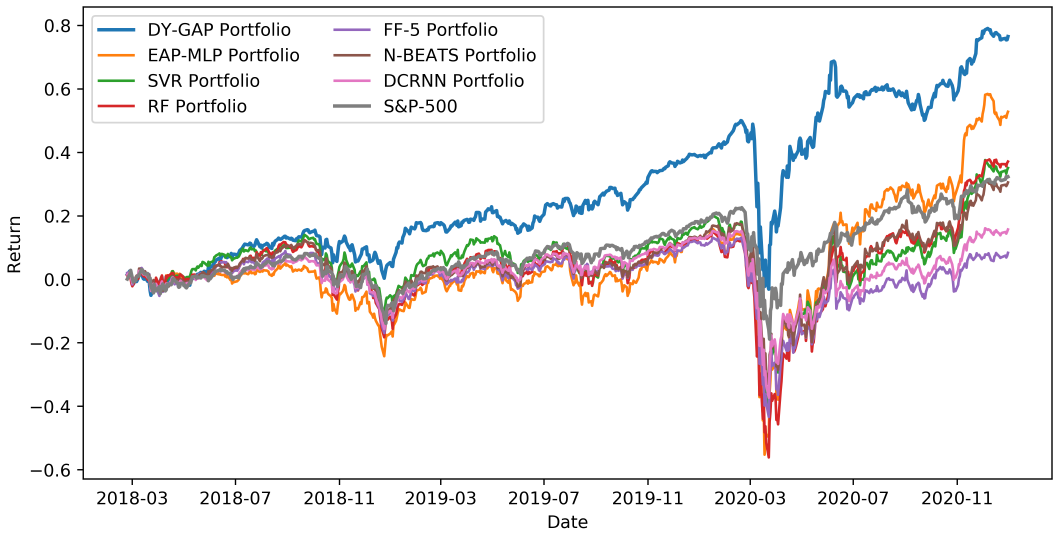


Figure 6: The portfolio performance from the daily data. For each trading day, we take long position at the top 10% of highest predicted stocks. The thick blue line is the cumulative return of our DY-GAP model in the test period.

Sharpe Ratio of our model, DY-GAP, is the highest with 1.18 and significantly higher than that of the second-best model, N-BEATS, with 0.95.

We perform a portfolio analysis on the S&P-500 stocks using daily data as well. Figure 6 shows the cumulative return from the portfolio constructed based on the daily prediction. DY-GAP maintains superior performance over all other models throughout the test period. Our cumulative return is two times higher than the S&P-500 index return. There is an interesting pattern that the up-down trend in the portfolio performance from all models echos the market return. All models, including the DY-GAP, are affected by the market's downside, particularly for unexpected events like the COVID-19 pandemic. However, the incorporation of network and spatial dependency allows DY-GAP to select the stock groups that generate relatively high returns in the middle of the pandemic. EAP-MLP is the second-best, especially in the latter part of 2020.

To examine the statistical significance, we conduct mean difference tests of portfolio return between DY-GAP and a given benchmark in Table 4. For example, the monthly historical return from DY-GAP is 1.5500% and that from MR is 1.0046%. 1.0046*** indicates

Table 4: Portfolio Performance

	Monthly Data			Daily Data		
	Average (%)	STD (%)	Sharpe Ratio	Average (%)	STD (%)	Sharpe Ratio
MR	1.0046***	4.0400	0.8614	0.0120***	1.7669	0.1081
FF-5	1.0980**	4.4340	0.8578	0.0116***	1.7450	0.1055
EAP-MLP	1.1455*	5.1880	0.7649	0.0733	2.2850	0.5092
ARIMA	1.1840*	4.9540	0.8279	0.0401*	1.6876	0.3772
SVR	1.0323***	4.4024	0.8123	0.0487*	1.8575	0.4165
Random Forest	1.0641**	5.6452	0.6529	0.0515*	2.0296	0.4032
FC-LSTM	1.1790*	4.8877	0.8356	0.0322**	2.0711	0.2468
N-BEATS	1.1046**	4.0400	0.9471	0.0424*	1.8450	0.3648
DCRNN	0.9121***	5.1650	0.6117	0.0218***	1.6104	0.2149
GAT	1.0820**	4.7730	0.7853	0.0220***	1.7669	0.1980
S&P-500	0.8236***	3.2534	0.8769	0.0449*	1.4892	0.4785
DY-GAP	1.5500	4.5617	1.1771	0.0854	1.9500	0.6955

Note: This Table reports the performance of long portfolios created based on the prediction of our proposed model (DY-GAP) and other benchmarks. At the beginning of each month (day), we hold the top 10% stocks based on the prediction of each model and at the end of the month (day), we liquidate all stocks. ***, **, * indicate the historical average return of DY-GAP is higher than a given benchmark at the 1%, 5%, and 10% significance level, respectively.

that DY-GAP return on average is significantly higher than MR at one percent level. In almost all cases, the daily and monthly portfolio constructed based on DY-GAP generates significantly higher returns with just one exception. This finding is consistent with the return prediction exercise and confirms the high forecasting accuracy of our model.

In addition, following Jiang et al. (2020) and Li & Rossi (2020), we test our proposed model performance by creating portfolios based on the model out-of-sample forecast. For each month (day) t, DY-GAP predicts the performance of each stock based on the information at $t-1$. We then sort stocks according to their predicted returns from low (decile 1) to high (decile 10), form ten equally weighted portfolios, and hold them for one month (day). Low (High) represents the decile portfolio that contains stocks that are expected to perform the worst (best). "High-Low" denotes the long-short portfolio that longs stocks with the highest expected returns and shorts stocks with the lowest expected returns.

Table 5: Portfolio Return on DY-GAP Predicted Sort

	Monthly		Daily	
	Return	t-stat	Return	t-stat
Low	1.0364	0.4675	-11.8063^{**}	-1.6717
2	4.1432	0.9863	-9.3565^{*}	-1.5058
3	8.1864^{*}	1.4534	5.3051	0.3175
4	10.0680^{*}	1.6300	5.5978	0.3613
5	11.4972^{**}	2.1344	15.9234^{**}	2.0584
6	10.3176^{**}	1.8048	17.4572^{**}	2.1589
7	10.5780^{**}	1.7525	18.5066^{***}	2.4818
8	13.9844^{***}	2.3873	17.8745^{***}	2.2536
9	16.0248^{***}	2.9738	21.3467^{***}	2.5356
High	18.2300^{***}	3.4826	22.2098^{***}	2.8850
High - Low	17.1936^{***}	3.4113	34.0161^{***}	3.8415

Note: This table presents the excess return for portfolios generated based on the DY-GAP model prediction. For each month (day) t, DY-GAP predicts the performance of each stock based on the information on month (day) $t-1$. We then sort stocks according to their predicted returns, form equally weighted decile portfolios, and hold them for one month (day). Low (High) represents the decile portfolio with stocks expected to perform the worst (best). "High-Low" denotes the long-short portfolio that goes long in the portfolio of stocks with the highest expected returns and short in the portfolio with the lowest expected returns. We report annualized excess returns together with their t-statistics. ***, **, * indicate significant at the 1%, 5% and 10% significance level, respectively.

Table 5 reports the average annualized excess returns along with the t-statistics. The reported values in the table are only from out-of-sample periods, e.g., for monthly data from January 2014 to December 2019 and for daily data from January 01-2018 to December 31-2020. Columns two and three report returns from one month holding period strategy on monthly data, and columns three and four reports one day holding period on daily data. The portfolio return increase almost monotonically from decile 1 to 10. A long-only strategy based on stocks in decile 10 earns an average excess return of 18% annually on the monthly holding period and of 22% annually on the daily holding period. The long-short portfolio strategies earn an annualized return of 17% and 34% on monthly and daily data, respectively. These

returns on long-short portfolios are statistically different from zero, signifying DY-GAP's ability to forecast stock returns accurately.

5.3. Fama-MacBeth Regressions Controlling for other Characteristics Predictors

In this section, we perform a Fama–MacBeth regression to test if our DY-GAP forecasts encompass and are not subsumed by the other established predictors in the literature (Li & Rossi, 2020). We implement the following Fama-MacBeth regressions with the DY-GAP model forecast after controlling for market, size, value, and momentum:

$$y_{i,t} = \alpha + \beta_1 \widehat{DYGAP}_{i,t|t-1} + cM_{i,t-1} + \epsilon_{i,t}, \quad for\ t = 1, ..., T \tag{14}$$

where, $y_{i,t}$ is stock's i realized excess return in month t ; $\widehat{DYGAP}_{i,t|t-1}$ is DY-GAP prediction for stock i's excess return on month t, conditional on information available as of time $t-1$; and $M_{j,t-1}$ is a vector of regressors controlling for market, size, value, and momentum risk factor for each firm at time $t-1$. For monthly data, we estimate changing β's of each factor for each firm using a 60-month rolling window. For daily data, we estimate changing β's of each factor for each firm using a 252-trading days rolling window.

The regression results are reported in Table 6. Columns two and three reports results from the monthly data and columns four and five report the results from daily data. In "w/FF3", we control for Fama-French three factors – Market, Size, and Value, – and we add Momentum factor in "w/FF4". Consistent with the hypothesis that our predicted returns contain information that is not subsumed by other characteristics based predictors reported in the literature, we find that the β_1 coefficient is positive, statistically significant and very stable across specifications - "w/FF3" and "w/FF4". Furthermore, we find that the coefficient of other common return predictors such as Market, Size, Value, and Momentum are very small compared to the coefficient of DY-GAP. Overall, the test results of this subsection indicate that the predictive power of DY-GAP cannot be fully explained by the existing pricing factors that have been widely used in the literature.

Table 6: Fama-MacBeth Regression on Next Month/Day's Excess Return

	Monthly		Daily	
Factors	w/FF3	w/FF4	w/FF3	w/FF4
DY-GAP	0.1371***	0.1444***	0.0527***	0.0527***
	(0.0150)	(0.0151)	(0.0053)	(0.0061)
Market	−0.0086***	−0.0088***	−0.0006***	−0.0007***
	(0.0011)	(0.0013)	(0.0001)	(0.0000)
SMB	0.0035***	0.0026***	0.0004***	0.0005***
	(0.0011)	(0.0010)	(0.0000)	(0.0000)
HML	0.0005	0.0017	−0.0005	−0.0004
	(0.0014)	(0.0017)	(0.0007)	(0.0012)
Momentum		−0.0049***		0.0003***
		(0.0009)		(0.0000)
Constant	0.0072***	0.0081***	0.0008***	0.0009***
	(0.0015)	(0.0014)	(0.0000)	(0.0000)

Note: This table reports the average value of Fama-MacBeth regressions. The dependent variable is the stocks' excess return in month t. Independent variables are the DY-GAP prediction for stocks excess return on month t, conditional on information available as of time $t-1$, β of Fama-French three factors (Market, SMB, and HML) in "w/FF3" and Fama-French three factors with Momentum in "w/FF4". Columns two and three report results from monthly data, and columns four and five report results from daily data. Newey-West adjusted t-statistics are in parentheses. ***, **, * are significant at the 1%, 5%, and 10% significance level, respectively.

5.4. Ability to Learn Dynamic Graph

In Section 3.2.2 we delineate the graph (network) learning technique of our DY-GAP model. In this section, we evaluate the network learning capacity of our proposed DY-GAP model. From the S&P-500 stocks daily data, it is evident to see the dynamic evolution of network in the U.S. equity market. Figure 7 shows the distribution of positive and negative edges during the studied period. As expected, in the equity market, more firms are positively associated, including those large firms in the S&P-500 index. As a result, most of the time, among all possible connections, the percentage of positive edges is 15%-30%, and the percentage of negative edges is 5%-12%. The remaining percentages are the insignificant

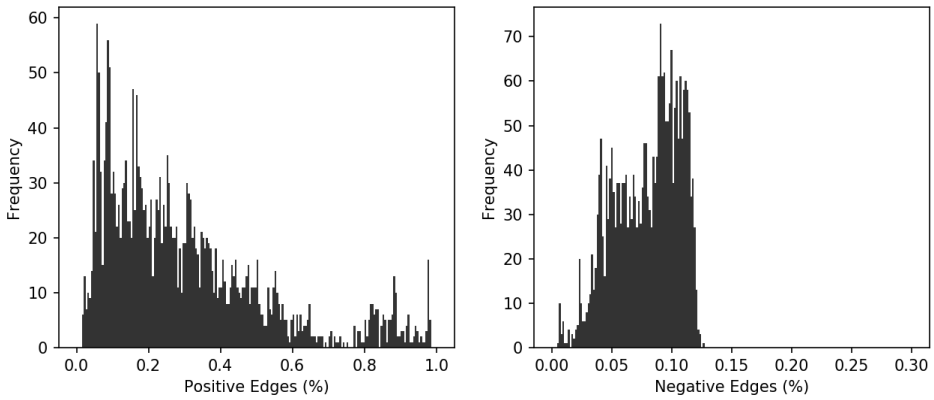


Figure 7: Histogram of the learned edges from S&P-500 daily data.

edges that are eliminated via the graph learning process. However, in some time periods, the percentage of positive edges reaches $> 90\%$. In these periods, most edges become significant, and our graph learning process accurately captures this time-varying pattern in the equity market network. Our in-depth analyses reveal that those periods with extremely high positive edges are associated with economic or political events. Figure 8 shows that, during the Eurozone crisis, Chinese stock market turbulence, Brexit, and Covid-19, the percentage of positive edges significantly increases. Our finding is consistent with the literature (Sandhu et al., 2016). During normal periods with a stable macro environment, a firm's performance might be more dependent on its own financing, investment, and operation decisions, rather than the macro environment and it is natural to see both positive and negative linkages among firms. By contrast, during a major event, most firms respond accordingly in the same direction. In other words, there is a similar upward or downward trend in stock prices, and consequently, positive linkages increase. This pattern is consistent with contagion effects and systemic risks during recessions and crises.

Besides, the impacts of major events are not identical. We use two events as examples, one is the Brexit in 2016, and another is the Covid pandemic in 2020, and show how the market network changes before and after those two events in Figure 9. Figures 9a-9e represent the market structure before, during, and after Brexit from June 2016 to October 2016. Figures

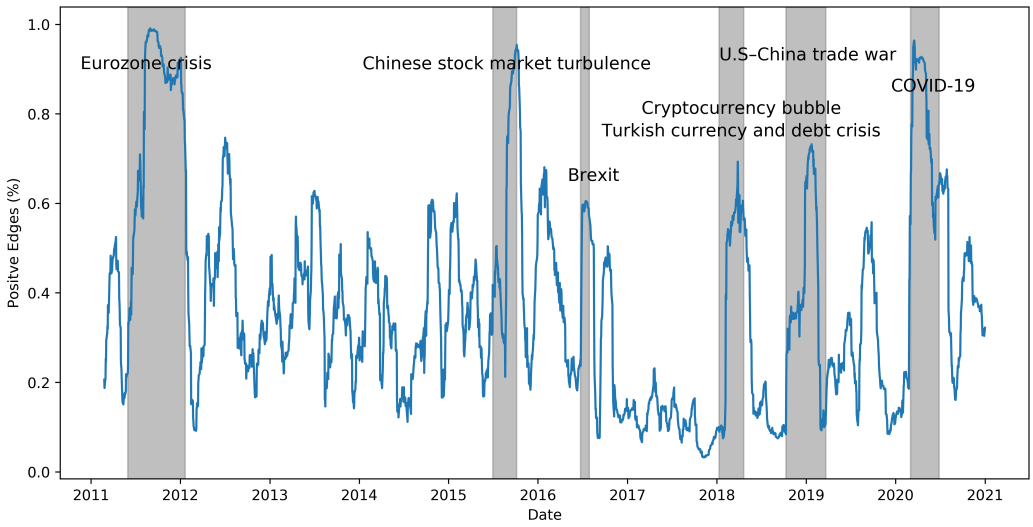
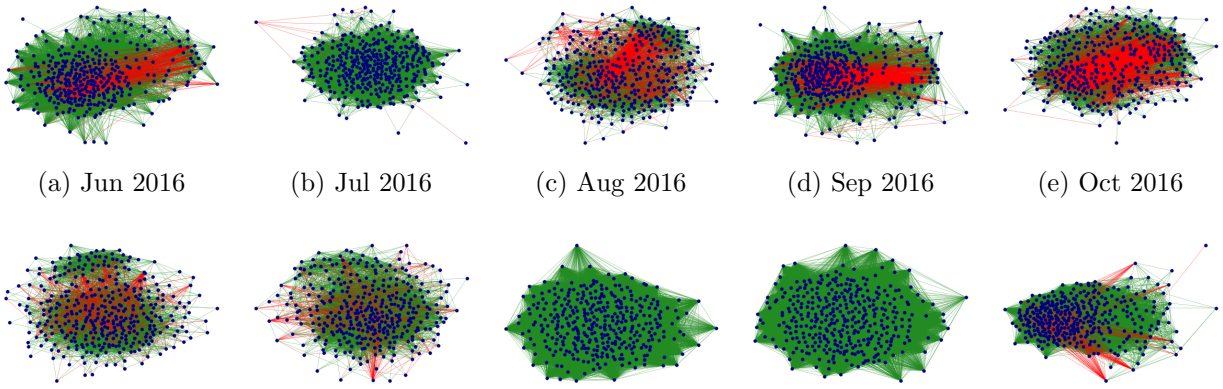


Figure 8: The learned positive edges from S&P-500 daily data. There are visible spikes in the degree of positive edges during significant financial, political, and economic events.



(a) Jun 2016 (b) Jul 2016 (c) Aug 2016 (d) Sep 2016 (e) Oct 2016

(f) Jan 2020 (g) Feb 2020 (h) Mar 2020 (i) Apr 2020 (j) May 2020

Figure 9: The network structure of S&P-500 stocks at different point of time. Figure 8a-8e are the market network before, during, and after the Brexit, and Figure 8f-8j are for the Covid pandemic. Green represents positive edges and red represents negative edges. During the normal economic condition, the market consists of a mixture of positive and negative edges. However, during the market shocks, the network is dominated by positive edges, representing the common effect across most firms.

9f-9j are for the Covid-19 pandemic impact on the market network structure from January 2020 to May 2020. During the Brexit, the network structure changed (sub-Figure 9b), but all the negative edges did not vanish overnight. The market also recovered quickly and went back to its normal state. However, during the outbreak of Covid-19 pandemic, almost all

edges turned to be positive (sub-Figures 9h and 9i). At the beginning of May, the market started to cope with the pandemic's initial setback and recover to a certain extent. In sub-Figure 9j, it is also noticeable that instead of scattered negative edges through the market, there exist a handful of companies that had negative edges. Further investigation reveals the good performance of this handful of companies, such as Amazon and Walmart, attributed to the booming online business during the peak months of the Covid-19 Pandemic.

6. Robustness Tests

In this section, multiple tests are performed to ensure the robustness and reliability of our proposed DY-GAP model, including different data structures, the sensitivity of the algorithm, and the contribution of different model components. First, in the ablation study, we evaluate the contribution of different model components. Second, we test the sensitivity of our model with different subsamples based on firm characteristics, data distribution and random selection. Third, we created a synthetic data with multiple time series features and compare our model performance against the benchmarks on the synthetic data. Lastly, we examine the flexibility of our model by incorporating macroeconomic factors into our model.

6.1. Ablation Study

To gain a better understanding of the contributions of individual components of our model, we perform an ablation study with five alternative versions of our model: (i) the proposed model with a signed network structure, attention mechanism, and diffusion convolution (DY-GAP), (ii) without any pre-existing correlation network structure (DY-GAT), (iii) considering both positive and negative relationships in a single network (DY-GAP-Unsigned), (iv) replacing diffusion convolution with simple graph convolution (GCN) as suggested in Kipf & Welling (2017) (DY-GCN), and (v) with only a single diffusion layer (DY-GAP-Single Diffusion).

Table 7 reports the prediction accuracy of alternative models. DY-GAT ignores the constraints on network connectivity and allows all firms to attend to all other firms. The

Table 7: Prediction Results from Ablation Study

	Monthly Data		Daily Data	
	RMSE	MAPE (%)	RMSE	MAPE (%)
DY-GAP	0.0853	52.2400	0.0233	15.6098
DY-GAT	0.1121	61.7160	0.0294	19.2406
DY-GAP-Unsigned	0.1072	58.4270	0.0261	17.0153
DY-GCN	0.0936	53.1476	0.0270	20.5101
DY-GAP-Single Diffusion	0.0908	52.6881	0.0240	16.9219

Note: This table reports the return prediction results from alternative versions of our model. DY-GAP is our proposed model, DY-GAT is a model without any network structure, DY-GAP-Unsigned is a model with unsigned networks using absolute values, DY-GCN is a model with simple graph convolution instead of diffusion convolution, and DY-GAP-Single Diffusion is a model with only a single diffusion layer. A lower value of RMSE and MAPE indicates better prediction performance.

change to the model leads to the performance degradation by 31% in monthly data and 26% in daily data compared to the model abiding by the restrictions. In DY-GAP-Unsigned, we provide a binary network structure for masked attention by ignoring the sign. This version does not differentiate between positive and negative connections and provides the same attention to both. As the result of modeling contradictory information identically, the model achieves a sub-optimal performance. For DY-GCN, we use the simple graph convolution neural network instead of the diffusion convolutional recurrent neural network. Simple convolution only models spatial dependency among firms, ignoring any temporal dependency. Earlier studies suggest that the historical average excess return be an important predictor for future excess return. Therefore, ignoring the critical temporal relation will result in sub-optimal model performance. The performance drops in DY-GCN is 10% in monthly data and 16% in daily data. In DY-GAP-Single Diffusion, we use only one diffusion layer and predict $\hat{Y}$ from initial $\mathcal{H}$, skipping the second layer of concatenating the return embedding H. Its performance decreases by 6% for monthly data and 3% for daily data. The rationale for this observation is that the concatenation of return embedding allows the model to incorporate historical information directly into the learning process. Although we learn

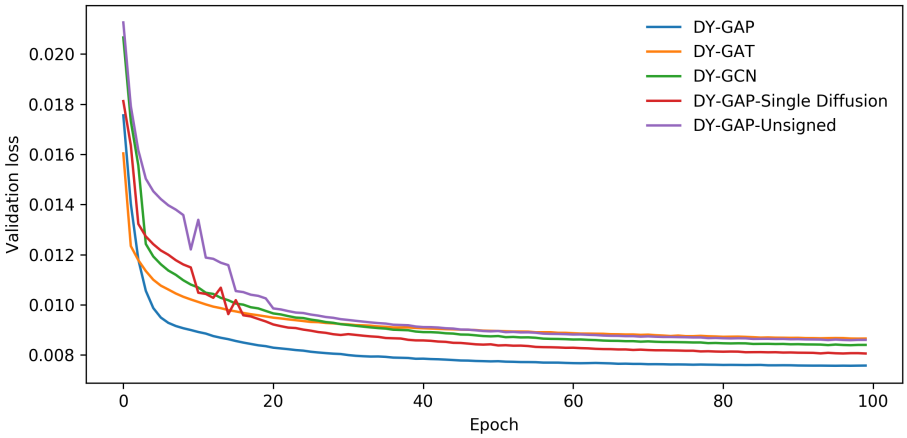


Figure 10: The loss curves of our proposed model and some other alternative of network and convolution operation. DY-GAP with both positive and negative network and diffusion convolution achieves the lowest validation error.

initial graphs from the return embedding, the graph learning process loses some information through masking operation and vanishing gradient. The concatenation function and second diffusion layers help us recover the lost information and thereby improve model performance.

Figure 10 shows the validation loss curves among all ablation models from the monthly data. The validation error of all the models converges in approximately 40 epochs. DY-GAP has the lowest validation error among all alternative versions. Both unsigned (DY-GAP-Unsigned) and the single diffusion model (DY-GAP-Single Diffusion) suffer sporadic gradient descents initially but eventually converge. The plausible reason for this slow convergence is that these two models access less information than others. In the unsigned graph, the machine learning model considers both positive and negative edges the same; therefore, learning meaningful representation becomes difficult. For single diffusion, we are only learning spatial-temporal dependency on the features. The diffusion convolution layer has no direct access to the historical return data; as a result, the learning process is slower.

6.2. Sensitivity Study

The ablation study in Section 6.1 provides the sensibility tests of model components and hyper-parameter optimizations. Here, we further perform a detailed sensitivity analysis of our model concerning different data subsamples. We divide all stocks into groups based on

different properties, including data distribution, skewness, volatility, return structure, and size.

We first divide both monthly and daily data sets according to their return distribution. Previous studies show that the distribution of financial returns exhibits "leptokurtic" – peaked in the center and has a fat tail (Mandelbrot, 1967; Fama, 1965). The stocks vary in the "fatness" of their return distribution. This prompts us to evaluate the prediction performance of our model in stocks with different distributions, especially when stock returns have a fat tail. Prior studies document that the tail index of stock returns lies in the range of 2.5 to 5.0 (Lux, 1996; Loretan & Phillips, 1994; Jansen & De Vries, 1991).[7]. The average tail index α is 3.14 and 3.5 for S&P500 and Russel 2000, respectively. We then compare the model between the firms with the tail index above and below 3.5 in both data sets. Table 8 presents the results. There is a significant difference in model accuracy between the two tail index groups in both daily and monthly data sets. The model achieves better accuracy for stocks with a lower tail index < 3.5 in all three performance matrices. In monthly data, lower tail index stocks outperform higher tail index stocks by 21%, and for daily data, it outperforms by 13%. The underperformance of our model for fatter tail stocks is expected because these stocks' returns are highly dispersed, have a large interquartile range (monthly 0.139 and daily 0.02), and high variance (monthly 0.0113 and daily 0.0196). It is relatively easier to train a model and learn its parameters for data with a small range and low variance (Kantz & Schreiber, 2004).

Second, we test the model performance in negatively and positively skewed stocks. On average, consistent with common literature (Lux, 1996), we observe a moderate right skewness (0.20) on monthly data, and on daily data, we observe a moderate left skewness (-0.17). We use the bottom and top 15% of stocks and run the model. The model performs better for left-skewed stocks. The difference is more significant for monthly data (RMSE 30%) than

[7]The tail index α is the rate of tail decay. The tail index measures the fatness of the tails (Haas & Pigorsch, 2009; Hill, 1975) Following(Hill, 1975), we use the Hill estimator to calculate tail index α.

Table 8: Sensitivity Analysis

	Monthly Data			Daily Data		
Group	RMSE	MAE	MAPE (%)	RMSE	MAE	MAPE (%)
Tail Index ≥ 3.5	0.0919	0.0709	55.0712	0.0224	0.0168	16.7942
Tail Index < 3.5	0.0640	0.0528	51.6870	0.0196	0.0143	14.2725
Negatively Skewed	0.0641	0.0486	44.8418	0.0192	0.0138	13.8281
Positively Skewed	0.0923	0.0681	53.4210	0.0203	0.0147	14.7243
High Volatility	0.1007	0.0742	58.0103	0.0250	0.0187	17.1220
Low Volatility	0.0645	0.0496	46.6504	0.0196	0.0143	13.2873
High Return	0.0808	0.0617	53.0959	0.0217	0.0159	15.0365
Low Return	0.0831	0.0606	52.6336	0.0235	0.0162	15.8396
Big Size	0.0805	0.0605	52.7416	0.0209	0.0155	15.5003
Small Size	0.0845	0.0685	53.7218	0.0200	0.0144	14.3602

Note: This table reports the prediction results of the proposed model (DY-GAP) for both S&P-500 stocks (monthly data) and Russell-3000 stocks (daily data) with different subsamples. A lower value of RMSE, MAE and MAPE indicates better performance. For Volatility, Return and Size, stocks are divided into tercile and only top and bottom 1/3 are reported.

daily data (RMSE 5%).

Thirdly, we group our stocks based on their return volatility. We divide stocks into three groups: high volatility, moderate volatility, and low volatility. Results show that the model performs relatively well for the low return volatility stocks than highly volatile stocks. This finding aligns with the general expectation of a machine learning model to perform better when the target value is less volatile, as most machine learning models struggle to find the appropriate parameters for highly volatile data (Kantz & Schreiber, 2004). In addition, we analyze our model performance in groups based on stock returns and firm size. For both data sets, we construct three return groups based on their average return over time – high return, moderate return, and low return – and three size groups based on their total assets – big, medium, and small. We separately test our model in each of these size groups and

reveal no significant difference in the model performance, shown in the last two panels in Table 8.

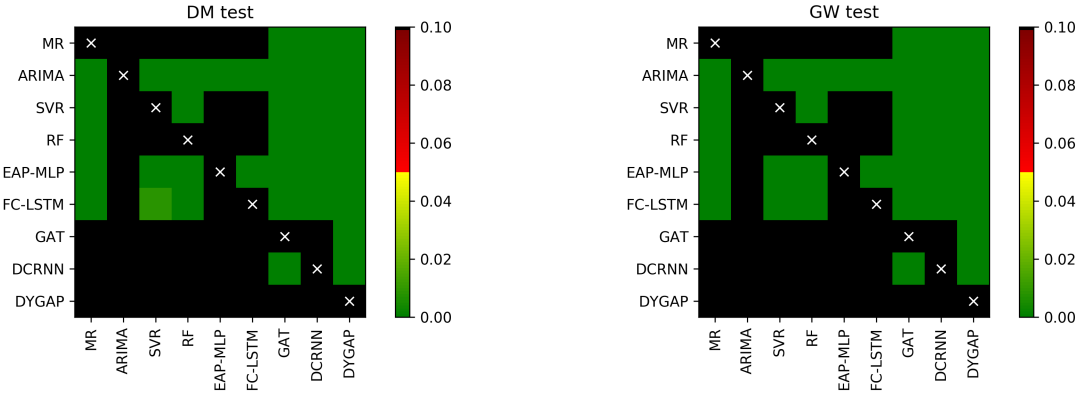
Although the prediction performance of our model varies across different groups, the difference is not very large. For example, in the second column in Table 8, the range of MAE among ten groups is only 0.0256. Even the lowest performing group "high volatility" still outperforms MR, FF-5, ARIMA, N-BEATS, and GAT for monthly data, and all benchmarks except DCRNN in daily data, as reported in Table 3. This suggests that our model is able to handle different data structures and can be applied to most firms. Overall, these results confirm the stability and robustness of the proposed DY-GAP in return prediction.

In addition to testing with different data structures, we further test the proposed DY-GAP sensitivity with different window Size K and randomly subsampling a portion of our original data. To test the model robustness on window size K for monthly data we use alternative windows of $K = 12, 24, 36$ and 48, and for daily data we use $K = 44, 66, 122$, and 252. Finally, to test model robustness on data attributes, we randomly select 100 stocks with replacements from our original data and run our DY-GAP model. We repeated the exercise ten times. The results from these experiments show that the prediction errors of our model are very similar, with different window sizes and different data samples. This indicates the consistency and reliability of our model. The detailed results of these two robustness tests are provided in Appendix B.

6.3. Synthetic Data

In this section, we create synthetic datasets containing the major characteristics of economics and financial time series, test our proposed model along with the benchmarks, and examine the model's robustness under given experiments. The synthetic data is constructed by first initializing it with two random time series and then adding the following factors:

- The GDP factor with the actual Gross Domestic Product (GDP) of five countries, including the USA, the UK, Japan, China, and Italy.



(a) Diebold-Mariano Test

(b) Giacomini-White Test

Figure 11: p-values from Diebold-Mariano Test and Giacomini-White Test. For both test, rejecting the null hypothesis indicates that the predictive ability of the model in horizontal axis is significantly more accurate than the model on vertical axis. The color spectrum represents the p-value of the t-stat. Black (p-value > 0.10) indicating that the null hypothesis is not rejected at 10% significant level, red ($0.05 <=$ p-value < 0.10) represent the null hypothesis is rejected at 10% significant level, the yellow to green (p-value < 0.05) represent the null hypotheses is rejected at 5% and lower significant level.

- Five seasonality and up-down trends in each time-series. We created one upward trending, one downward trending, one initially down and then upward trending, one with seasonality, and one with reversal trending. These seasonality and trend factors vary on wavelength, amplitude, and phase.

- Each country's actual industrial production time series.

- Holiday factor, with the multiplier values: 2.0 for Christmas Day and 1.2 for all other holidays.

- Weekday factor, a factor of 1.1 for Monday and 1.05 for Friday.

In total, there are fifty-time series $5 \times 2 \times 5$. Each country with ten-time series consisting of two random components and five seasonality components.[8]

Table 9 presents the prediction performance results using synthetic data. We compare our model with all benchmarks in terms of three performance metrics RMSE, MAE, and

[8]The details code and the synthetic data will be shared in the published version for reproducibility.

Table 9: Result from Synthetic Data

	RMSE	MAE	MAPE (%)	t-stat (DM)	t-stat (GW)
MR	0.0640	0.0510	44.3047	5.4106***	25.87674***
ARIMA	0.0944	0.0749	60.7537	22.7274***	171.9241***
EAP-MLP	0.0800	0.0636	55.5862	20.7090***	164.8567***
FC-LSTM	0.0754	0.0595	49.2249	13.1950***	116.4760***
SVR	0.0731	0.0586	49.9343	10.7069***	79.5127***
Random Forest	0.0676	0.0538	46.8725	8.5404***	56.4640***
DCRNN	0.0651	0.0495	41.1416	5.4832***	48.5219***
GAT	0.0605	0.0482	41.9682	6.9749***	45.7709***
DY-GAP	0.0585	0.0465	40.5913	—	—

Note: The table reports the prediction results of Synthetic data. We created fifty unique time series with multiple random, macroeconomic, and financial time series components. The two t-statistics are Diebold-Mariano Test and Giacomini-White (GW) tests with the null hypothesis that the mean absolute forecast error of the "benchmark model" is smaller than or equal to that of "DY-GAP". Rejecting the null hypothesis indicates that the prediction of DY-GAP is significantly more accurate than a given benchmark. *** indicate that the t-stat is significant at 1% significance level.

MAPE, in columns one, two, and three, respectively. The proposed model DY-GAP outperforms others in all three performance metrics. For example, DY-GAP reduces RMSE and outperforms multi-layer perceptron by 26%, LSTM by 22%, support vector regression by 20%, and random forest by 13%. The test results in Table 9 already show that DY-GAP has the lowest forecasting error. We go one step further to determine whether this difference is significant or due to the specific choice of data values in the sample. For this purpose, we perform DM (Diebold & Mariano, 2002), and GW (Giacomini & White, 2006) tests and reported the t-statistics in columns four and five in Table 9. Here, we applied the one-sided version of the DM and GW tests. Rejecting the null hypothesis in both tests means that the forecasts of DY-GAP are significantly better than those of a given "benchmark model", e.g., MR, ARIMA, SVR. Table 9 shows that both the null hypotheses are rejected at the 1% significant level against all benchmarks in both DM and GW tests. This signifies that with randomly created synthetic data, DY-GAP remains superior in prediction performance.

Table 10: Prediction with Macroeconomic Data

	RMSE	MAE	MAPE (%)	t-stat (DM)	t-stat (GW)
DY-GAP	0.0853	0.0632	52.2400	1.7859**	5.1132**
DY-GAP w/Macro	0.0824	0.0607	50.7569	—	—

Note: The table reports the prediction results of model with Macroeconomic data. The two t-statistics are Diebold-Mariano Test and Giacomini-White (GW) tests with the null hypothesis that the forecast error of the DY-GAP without macroeconomic data is smaller than or equal to that of DY-GAP model with macro economic data. ***, **, * indicate to reject the null hypothesis at the 1%, 5%, and 10% significance level, respectively.

Figure 11 depicts the heatmap of the p-values from the DM and GW tests for all pairwise model comparisons among our proposed DY-GAP and other benchmarks. Here, we test the null hypothesis that the forecasting errors of model A on the vertical axis are worse with larger forecast errors than the forecasting errors of model B on the horizontal axis. Black represents the null hypothesis is not rejected even at 10% significance level, red represents the null hypothesis is rejected at 10% significance level, and the green represents the null hypothesis is rejected at 5% or lower significance level. The figure clearly highlights that the three most sophisticated models, GAT, DCRNN, and DY-GAP, are also the top three best-performing models. The prediction accuracy of these three models is significantly better than other traditional models. Among the traditional models, multiple regression outperforms ARIMA, SVR, RF, EAP-MLP, and LSTM.

6.4. Macroeconomic Factor

Our DY-GAP framework provide the flexibility to incorporating more factors along with firm characteristics. To investigate whether more factors can improve the prediction accuracy of our model, we consider macroeconomic factors. As firm performance can be affected by the macro environment and therefore macro factors can be inputs in equity valuation. To do so, we can simply extend the dimension of the model input X in Figure 3.

Multiple macroeconomic factors are recognized and widely used as a factors in asset pricing literature (Gu et al., 2020b). In this analysis we incorporated Gross Domestic Product

(GDP), Personal Consumption Expenditures (PCE), Industrial Production Total Index (IP), Unemployment Rate, Advance Retail Sales, Producer Price Index (PPI), Total Construction Spending, Federal Fund Effective Rates, Interest Rates, Moody's Seasoned AAA Corporate Bond Yield, Chicago Board Options Exchange's Volatility Index (VIX), S&P-500 index return (S&P), and Economic Policy Uncertainty Index (EPU).[9]

In Table 10, our experiments show that the incorporation of macroeconomic information in the monthly model slightly increases the prediction accuracy of our model. Recall that the prediction errors (RMSE, MAE, and MAPE) of DY-GAP without macroeconomic variables are 0.0853, 0.0632, and 52.2400, respectively. With macro factors, those prediction errors reduce to 0.0824, 0.0607, and 50.7569, an improvement of 3%, 4%, and 3%, respectively. The DM and GW tests on the forecasting errors of the model with and without macroeconomic indicators show that the improvement is significant at a 5% significance level. Results suggest that the improvement in prediction from the incorporation of macro factors is statistically significant but economically minor. One plausible explanation is that those macro factors are at least partially captured by firms' characteristics and learned networks (as demonstrated in Section 5.4). Therefore, a large portion of the information content of these macroeconomic factors is already included in our model.

7. Conclusion

In this paper, we propose a graph neural network-based approach for asset return prediction. This work offers a novel framework to incorporate network information into asset pricing by capturing the dynamic inter-connection among firms and integrating firm characteristics.

Our model outperforms multiple well-known asset pricing and advanced machine learning models in terms of prediction accuracy and portfolio performance. This outperformance

[9]The data for S&P-500 and VIX are obtained from Yahoo Finance; EPU is downloaded from the economic policy uncertainty website: `https://www.policyuncertainty.com/`; and all other macroeconomics indicators are extracted from the Federal Reserve Bank of St. Louis.

is persistent among different groups and robust under various conditions. We also conduct prediction analysis with several alternative models in the ablation study and show that DY-GAP with the positive and negative correlation networks and diffusion convolution layer performs the best. Our findings confirm that the firm interconnection is bidirectional and relevant to the market analysis. The positive and negative relations must be treated differently because they serve different roles in the market. Incorporating an effective network representation into the model for spatial and temporal relations enhances the stock price prediction and, more profoundly, improves our understanding of the network structure in the financial market.

Our model sheds light on future studies. For example, we learn latent networks from firm characteristics and return embedding while overlooking some previously used static networks, such as sector similarity and supply-chain networks. What happens if we learn in multi-type heterogeneous networks? Does the performance improve with rich information or decline from the noisy and even conflicting content of heterogeneous data? Can we replace the complex recurrent diffusion convolution layer with any other simple recurrent neural network without affecting the performance? In addition, because of the complex nature of the model, we only test the model on the firms of two index constituents. Further investigation is needed to vet its superiority in the whole equity market. Therefore, like any other machine learning-based model, we suggest users exercise caution and good judgment before applying the technique in real-world trading.

References

Acemoglu, D., Ozdaglar, A., & Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, *105*, 564–608.

Adrian, T., Crump, R. K., & Moench, E. (2015). Regression-based estimation of dynamic asset pricing models. *Journal of Financial Economics*, *118*, 211–244.

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike* (pp. 199–213). Springer.

Assaf, A., Bilgin, M. H., & Demir, E. (2022). Using transfer entropy to measure information flows between cryptocurrencies. *Physica A: Statistical Mechanics and its Applications*, *586*, 126484.

Atwood, J., & Towsley, D. (2016). Diffusion-convolutional neural networks. In *Advances in neural information processing systems* (pp. 1993–2001).

Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. In *3rd International Conference on Learning Representations, ICLR 2015*.

Bikbov, R., & Chernov, M. (2010). No-arbitrage macroeconomic determinants of the yield curve. *Journal of Econometrics*, *159*, 166–182.

Billio, M., Caporin, M., Panzica, R., & Pelizzon, L. (2016). The impact of network connectivity on factor exposures, asset pricing and portfolio diversification, .

Borji, A., & Itti, L. (2013). State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*, 185–207.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Box, G. E., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*, *65*, 1509–1526.

Bruynseels, L., & Cardinaels, E. (2014). The audit committee: Management watchdog or personal friend of the ceo? *The accounting review*, *89*, 113–145.

Buraschi, A., & Porchia, P. (2012). Dynamic networks and asset pricing. In *AFA 2013 San Diego Meetings Paper*.

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, *21*, 1509–1531.

Cao, D., Wang, Y., Duan, J., Zhang, C., Zhu, X., Huang, C., Tong, Y., Xu, B., Bai, J., Tong, J. et al. (2020). Spectral temporal graph neural network for multivariate time-series forecasting. *Advances in Neural Information Processing Systems*, *33*.

Carvalho, V. M. (2014). From micro to macro via production networks. *Journal of Economic Perspectives*, *28*, 23–48.

Chen, L., Pelger, M., & Zhu, J. (2020). Deep learning in asset pricing. *Available at SSRN 3350138*, .

Cho, K., Courville, A., & Bengio, Y. (2015). Describing multimedia content using attention-based encoder-decoder networks. *IEEE Transactions on Multimedia*, *17*, 1875–1886. doi:10.1109/TMM.2015.2477044.

Chollet, F. et al. (2015). Keras. https://keras.io. URL: `https://keras.io`.

Chordia, T., & Shivakumar, L. (2002). Momentum, business cycle, and time-varying expected returns. *The journal of Finance*, *57*, 985–1019.

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. In *NIPS 2014 Workshop on Deep Learning*.

Cohen, L., & Frazzini, A. (2008). Economic links and predictable returns. *The Journal of Finance*, *63*, 1977–2011.

Defferrard, M., Bresson, X., & Vandergheynst, P. (2016). Convolutional neural networks on graphs with fast localized spectral filtering. *Advances in neural information processing systems*, *29*, 3844–3852.

Di Cerbo, L. F., & Taylor, S. (2020). Graph theoretical representations of equity indices and their centrality measures. *Quantitative Finance*, (pp. 1–15).

Diao, Z., Wang, X., Zhang, D., Liu, Y., Xie, K., & He, S. (2019). Dynamic spatial-temporal graph convolutional neural networks for traffic forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 890–897). volume 33.

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, *20*, 134–144.

Diebold, F. X., & Yılmaz, K. (2015). *Financial and macroeconomic connectedness: A network approach to measurement and monitoring*. Oxford University Press, USA.

Dimpfl, T., & Peter, F. J. (2013). Using transfer entropy to measure information flows between financial markets. *Studies in Nonlinear Dynamics and Econometrics*, *17*, 85–102.

Dong, H., Guo, X., Reichgelt, H., & Hu, R. (2020). Predictive power of arima models in forecasting equity returns: a sliding window method. *Journal of Asset Management*, *21*, 549–566.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector regression machines. *Advances in Neural Information Processing Systems*, *9*, 155–161.

Elliott, M., Golub, B., & Jackson, M. O. (2014). Financial networks and contagion. *American Economic Review*, *104*, 3115–53.

Esmailian, P., & Jalili, M. (2015). Community detection in signed networks: the role of negative ties in different scales. *Scientific Reports*, *5*, 1–17.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business*, *38*, 34–105.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, *25*, 23–49.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*, 1–22.

Fang, S., Zhang, Q., Meng, G., Xiang, S., & Pan, C. (2019). Gstnet: Global spatial-temporal network for traffic flow prediction. In *IJCAI* (pp. 2286–2293).

Feng, F., He, X., Wang, X., Luo, C., Liu, Y., & Chua, T.-S. (2019). Temporal relational ranking for stock prediction. *ACM Transactions on Information Systems (TOIS)*, *37*, 1–30.

Feng, G., Giglio, S., & Xiu, D. (2020). Taming the factor zoo: A test of new factors. *The Journal of Finance*, *75*, 1327–1370.

Geng, X., Li, Y., Wang, L., Zhang, L., Yang, Q., Ye, J., & Liu, Y. (2019). Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting. In *Proceedings of the AAAI conference on artificial intelligence* (pp. 3656–3663). volume 33.

Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, *74*, 1545–1578.

Glasserman, P., & Young, H. P. (2015). How likely is contagion in financial networks? *Journal of Banking & Finance*, *50*, 383–399.

Grover, A., & Leskovec, J. (2016). node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 855–864).

Grullon, G., Underwood, S., & Weston, J. P. (2014). Comovement and investment banking networks. *Journal of Financial Economics*, *113*, 73–89.

Gu, S., Kelly, B., & Xiu, D. (2020a). Autoencoder asset pricing models. *Journal of Econometrics*, .

Gu, S., Kelly, B., & Xiu, D. (2020b). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*, 2223–2273.

Haas, M., & Pigorsch, C. (2009). Financial economics, fat-tailed distributions. *Encyclopedia of Complexity and Systems Science*, *4*, 3404–3435.

Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: deep portfolios. *Applied Stochastic Models in Business and Industry*, *33*, 3–12.

Herskovic, B. (2018). Networks in production: Asset pricing implications. *The Journal of Finance*, *73*, 1785–1818.

Herskovic, B., Kelly, B., Lustig, H., & Van Nieuwerburgh, S. (2020). Firm volatility in granular networks. *Journal of Political Economy*, *128*, 4097–4162.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, (pp. 1163–1174).

Ho, T. K. (1995). Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition* (pp. 278–282). IEEE volume 1.

Hou, K. (2007). Industry information diffusion and the lead-lag effect in stock returns. *The Review of Financial Studies*, *20*, 1113–1138.

Huang, H., Yoo, S., Yu, D., & Qin, H. (2015). Density-aware clustering based on aggregated heat kernel and its transformation. *ACM Trans. Knowl. Discov. Data*, *9*, 29:1–29:35.

Hwang, B.-H., & Kim, S. (2009). It pays to have friends. *Journal of financial economics*, *93*, 138–158.

Ibrahim, A., Kashef, R., Li, M., Valencia, E., & Huang, E. (2020). Bitcoin network mechanics: Forecasting the btc closing price using vector auto-regression models based on endogenous and exogenous feature variables. *Journal of Risk and Financial Management*, *13*, 189.

Jansen, D. W., & De Vries, C. G. (1991). On the frequency of large stock returns: Putting booms and busts into perspective. *The review of economics and statistics*, (pp. 18–24).

Jiang, J., Kelly, B. T., & Xiu, D. (2020). (re-) imag (in) ing price trends. *Chicago Booth Research Paper*, .

Kantz, H., & Schreiber, T. (2004). *Nonlinear time series analysis* volume 7. Cambridge university press.

Kelly, B. T., Pruitt, S., & Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, *134*, 501–524.

Kim, R., So, C. H., Jeong, M., Lee, S., Kim, J., & Kang, J. (2019). Hats: A hierarchical graph attention network for stock movement prediction. *arXiv preprint arXiv:1908.07999*, .

Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017*.

Klicpera, J., Weißenberger, S., & Günnemann, S. (2019). Diffusion improves graph learning. *Advances in Neural Information Processing Systems*, *32*, 13354–13366.

Korbel, J., Jiang, X., & Zheng, B. (2019). Transfer entropy between communities in complex financial networks. *Entropy*, *21*, 1124.

Le, T. H., Do, H. X., Nguyen, D. K., & Sensoy, A. (2020). Covid-19 pandemic and tail-dependency networks of financial assets. *Finance research letters*, (p. 101800).

Li, B., & Rossi, A. G. (2020). Selecting mutual funds from the stocks they hold: A machine learning approach. *Available at SSRN 3737667*, .

Li, C., Song, D., & Tao, D. (2019). Multi-task recurrent neural networks and higher-order markov random fields for stock price movement prediction: Multi-task rnn and higer-order mrfs for stock price classification. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 1141–1151).

Li, W., Bao, R., Harimoto, K., Chen, D., Xu, J., & Su, Q. (2020). Modeling the stock relation with graph network for overnight stock movement prediction. In C. Bessiere (Ed.), *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20* (pp. 4541–4547).

Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2018). Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. In *International Conference on Learning Representations*.

Loretan, M., & Phillips, P. C. (1994). Testing the covariance stationarity of heavy-tailed time series: An overview of the theory with applications to several financial datasets. *Journal of empirical finance*, *1*, 211–248.

Lux, T. (1996). The stable paretian hypothesis and the frequency of large returns: an examination of major german stocks. *Applied financial economics*, *6*, 463–475.

Mandelbrot, B. (1967). The variation of some other speculative prices. *The Journal of Business*, *40*, 393–413.

Muslu, V., Rebello, M., & Xu, Y. (2014). Sell-side analyst research and stock comovement. *Journal of Accounting Research*, *52*, 911–954.

Namaki, A., Shirazi, A., Raei, R., & Jafari, G. (2011). Network analysis of a financial market based on genuine correlation and threshold method. *Physica A: Statistical Mechanics and its Applications*, *390*, 3835–3841.

Oreshkin, B. N., Carpov, D., Chapados, N., & Bengio, Y. (2019). N-beats: Neural basis expansion analysis for interpretable time series forecasting. In *International Conference on Learning Representations*.

Ozdagli, A., & Weber, M. (2017). *Monetary policy through production networks: Evidence from the stock market*. Technical Report National Bureau of Economic Research.

Pai, P.-F., & Lin, C.-S. (2005). A hybrid arima and support vector machines model in stock price forecasting. *Omega*, *33*, 497–505.

Perozzi, B., Al-Rfou, R., & Skiena, S. (2014). Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 701–710).

Post, O. (2012). *Spectral analysis on graph-like spaces* volume 2039. Springer Science & Business Media.

Rogers, L. C., & Veraart, L. A. (2013). Failure and rescue in an interbank network. *Management Science*, *59*, 882–898.

Sandhu, R. S., Georgiou, T. T., & Tannenbaum, A. R. (2016). Ricci curvature: An economic indicator for market fragility and systemic risk. *Science advances*, *2*, e1501495.

Sensoy, A., Ozturk, K., Hacihasanoglu, E., & Tabak, B. M. (2017). Not all emerging markets are the same: A classification approach with correlation based networks. *Journal of Financial Stability*, *33*, 163–186.

So, M. K., Chu, A. M., & Chan, T. W. (2020). Impacts of the covid-19 pandemic on financial market connectedness. *Finance Research Letters*, (p. 101864).

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, *27*, 3104–3112.

Tumminello, M., Lillo, F., & Mantegna, R. N. (2010). Correlation, hierarchies, and networks in financial markets. *Journal of economic behavior & organization*, *75*, 40–58.

Uddin, A., & Yu, D. (2020). Latent factor model for asset pricing. *Journal of Behavioral and Experimental Finance*, *27*, 100353.

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2018). Graph attention networks. In *International Conference on Learning Representations*.

Wang, C., & Xie, L. (2010). Information diffusion and overreaction: Evidence from the chinese stock market. *Emerging Markets Finance and Trade*, *46*, 80–100.

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, *21*, 1455–1508.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Philip, S. Y. (2020). A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, *32*, 4–24.

Yu, P., & Yan, X. (2020). Stock price prediction based on deep neural networks. *Neural Computing and Applications*, *32*, 1609–1628.

Zhang, L., Aggarwal, C., & Qi, G.-J. (2017). Stock price prediction via discovering multi-
1000 frequency trading patterns. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*.

Appendix A. Details of Baseline Models

ARIMA (Box et al., 2015): Auto-Regressive Integrated Moving Average (ARIMA) model is a popular baseline for time series prediction model. For ARIMA, we use data from 1990 January to December 2013 as the training set and the data from January 2014 to December 2019 as the testing set. The optimal lags for AR and MA are selected based on Akaike information criterion (Akaike, 1998) and Ljung–Box test (Box & Pierce, 1970).

Multivariate Regression (MR): We use a simple multivariate regression with all the fundamental features in our model. For the Linear Regression model, we use data from 1990 January to December 2013 as training data and from January 2014 to December 2019 as testing data.

FF5 (Fama & French, 2015): Fama French Five-Factor model is plausibly the most common benchmark for the asset pricing model. The model extend CAPM market factor by adding size, value, profitability and investment factors. The monthly beta for Fama French five-factor model is available at `https://wrds-www.wharton.upenn.edu/`. For each firm, the beta is calculated based on the last five years' historical data.

EAP-MLP (Gu et al., 2020b): Empirical Asset Pricing via Machine Learning (EAP-ML) is a seminal paper in the machine learning-based Asset Pricing model. In this work, the authors propose a MLP based framework with different hidden layers and neurons for firms' return prediction. We use the best model reported in the paper: three hidden layers with 32, 16, and 8 neurons, respectively.

Support Vector Regression (SVR) (Drucker et al., 1996): SVR is the extension of the Support Vector Machine (SVM) to predict the continuous output. It is a kernel-based machine learning algorithm that acknowledges the presence of non-linearity in the data. Instead of minimizing the error rate, SVR tries to fit the error within a certain threshold α. We use a nonlinear SVR with the radial basis function (RBF) kernel for our experiments.

Random Forest (RF) (Ho, 1995): Random forest is a tree-based ensemble learning algorithm for classification or regression. For regression tasks, it fits a number of decision trees

on various sub-samples of the dataset and uses the average of the prediction across decision trees. For our experiments, we use the scikit-learn standard version of Random Forest regression with a maximum depth of 20.

N-BEATS (Oreshkin et al., 2019): N-BEATS is a time series forecasting model based on deep neural network architecture with backward and forward link connections. We use the open-source code from `https://github.com/philipperemy/n-beats`. For our experiments, we set the backcast length 24, and the hidden layer unit is 32.

FC-LSTM (Sutskever et al., 2014): The LSTM framework for univariate time-series forecasting is proposed in (Sutskever et al., 2014). The source code is available at `https://github.com/farizrahman4u/seq2seq`. We use four recurrent LSTM layers with 56 hidden LSTM units in each layer. The model is trained with the same batch size as other models and MAE as the loss function. Early stopping is performed based on the model validation loss.

GAT (Veličković et al., 2018): Graph Attention Network (GAT) is a graph neural network architecture that uses the attention mechanism over neighborhoods' nodes features to assign different weights to different nodes. The model is mainly used for node classification. Because of the close similarity with our proposed model, we use GAT as a comparison model. We use the publicly available source code from `https://github.com/PetarV-/GAT` and make several modifications to cater to our purpose. In GAT we use both return and fundamentals together as features to learn the network. We use a two-layer GAT model with eight attention heads and sixteen output channels, followed by one dense layer with ReLu activation. These hyperparameters are selected on a grid search on the values reported in Table 2 using validation data.

DCRNN (Li et al., 2018): DCRNN is proposed for forecasting traffic by considering both spatial and temporal dependencies in the traffic flow. We use the publicly available source code from `https://github.com/liyaguang/DCRNN` with necessary modification for stock data. DCRNN requires a prior topology for which we use the correlation network. For the

horizon size, we use 36 months for monthly data and 126 days for daily data, which is the same as other models. For the hyperparameters similar to GAT we perform a grid search on the values reported in Table 2 and selected the best performing values on validation data.

Appendix B. Experimental Supplements

Appendix B.1. Sensitivity Test with Different Window Size

Table B.11 reports the sensitivity of the proposed DY-GAP model in relation to different window sizes (K) of the rolling window. Two important input parameters of DY-GAP depend on hyperparameter K. First, the historical return for the embedding layer, and second, the historical return to calculate the Pearson correlation for the masked attention. Panel A reports the result from monthly data. We use four different window sizes, 12 months, 24 months, 36 months, and 48 months. Panel B reports the daily data, and the four different window sizes are 44 trading days, 66 trading days, 122 trading days, and 252 trading days.

This sensitivity test demonstrates a weak positive relation between window size and performance. Especially when K is very small, i.e., 12 months for monthly and 44 trading days for daily, the performance is worst. However, the performance difference is negligible between the other three window sizes, i.e., 24, 36, and 48 months for monthly data and 66, 122, and 252 days for daily data. For daily data, the model performance peaks at 252 trading days. For monthly data, the performance peaks at 36 months and then slightly decline for 48 months. This indicates that increasing window size does not necessarily improve performance. With a longer window size, some of the recent trends may consider less influential in the graph learning process. In addition, a larger window size also increases computational complexity. Therefore, it is important to find the balance between the window size of the rolling window and model performance in relation to accuracy and complexity.

Appendix B.2. Sensitivity Test with Different Random Sub-samples

This appendix provide a further robustness test on the sensitivity of our model in relation to data. For both the daily and monthly data, we run our algorithm ten times on a random

Table B.11: Sensitivity Analysis with Different Window Size (K)

Window Size (K)	RMSE	MAE	MAPE (%)
Panel A: Monthly Data			
12 Months	0.1008	0.0838	56.6881
24 Months	0.0918	0.0638	53.0461
36 Months	0.0853	0.0632	52.2400
48 Month	0.0891	0.0677	52.6881
Panel B: Daily Data			
44 Days	0.0294	0.0179	17.0217
66 Days	0.0247	0.0167	15.7521
122 Days	0.0233	0.0161	15.6098
252 Days	0.0222	0.0157	14.7210

Note: This Table reports the return prediction results from our proposed model (DY-GAP) with different window size (K) of rolling window. The window size is used to learn latent embedding and calculate pearson correlation for masked attention. A lower value of RMSE, MAE and MAPE indicates better performance.

sample, where each time, we randomly select 100 stocks with replacements. This random sub-sampling allows us to select a smaller cluster of the data with different attributes. The prediction results of this sensitivity test are reported in Table B.12. Columns 1-3 represent the RMSE, MAE, and MAPE from monthly data, and columns 4-6 report the RMSE, MAE, and MAPE from the daily data. The last two rows report the average and standard deviation of ten rounds. For monthly data, the highest RMSE and MAE are in the fourth row with 0.0924 and 0.0708, respectively, and MAPE in the seventh row with 56.38. The average for all ten runs is 0.0843 for RMSE, 0.0645 for MAE, and 53.4730 for MAPE. The average is comparable yet slightly better than DY-GAP's original result on the full sample in Table 3.

In addition, we also construct the long-short portfolio based on the prediction results on each random seed on monthly data, shown in Figure B.12. The red line represents the cumulative weighted average returns of all the 100 stocks during the test period, and the green line represents the cumulative portfolio returns during the test period. For each

Table B.12: Sensitivity Analysis in Random Data Sub-Samples

	Monthly Data			Daily Data		
	RMSE	MAE	MAPE (%)	RMSE	MAE	MAPE (%)
1	0.0831	0.0601	52.7638	0.0247	0.0148	14.3869
2	0.0842	0.0637	52.9357	0.0278	0.0157	15.6636
3	0.0809	0.0678	51.9820	0.0273	0.0150	14.6031
4	0.0924	0.0708	55.8725	0.0246	0.0153	15.0020
5	0.0857	0.0632	52.4405	0.0249	0.0154	14.9824
6	0.0723	0.0605	51.1343	0.0240	0.0148	14.4950
7	0.0883	0.0667	56.3835	0.0255	0.0158	15.3837
8	0.0871	0.0644	53.2030	0.0252	0.0150	15.4830
9	0.0830	0.0623	53.7047	0.0247	0.0149	15.1460
10	0.0857	0.0650	54.3103	0.0242	0.0147	14.4210
Average	0.0843	0.0645	53.4730	0.0253	0.0151	14.9567
STD	0.0053	0.0033	1.6535	0.0013	0.0004	0.4652

Note: This Table reports the prediction results from the sensitivity study of the proposed model (DY-GAP). For both S&P-500 stocks (monthly data) and Russell-3000 stocks (daily data), we randomly select 100 stock in each time and run the model. In total, we run 10 random sampling with replacement. A lower value of RMSE, MAE and MAPE indicates better performance.

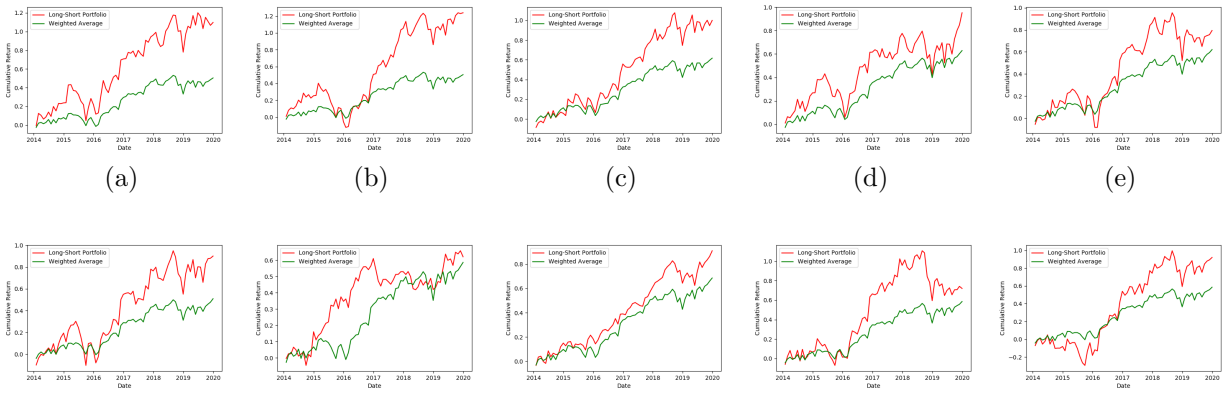


Figure B.12: Long-short portfolio with different seeds on monthly data

random seed, we take a long position on the ten stocks with the highest predicted returns and a short position on the ten stocks with the lowest predicted returns in the next time

period. We re-balance our portfolio at the end of each month. Figure B.12 shows that in almost all the random seeds, the long-short portfolio based on DY-GAP outperforms the simple weighted average by a significant margin.

Drivers and pass-through of the EU ETS price: evidence from the power sector

Yiyi Bai* Samuel J. Okullo†

01-October-2021

Abstract

Despite low and volatile allowance prices, power generators responded largely in line with theory to the European Union (EU) Emission Trading System (ETS). For the period between 2008 and 2018, we find that the qualitative response of the EU ETS allowance price to variations in the pre-eminent drivers of short-term fuel-switching, both the gas and coal price, is as economic theory prescribes. We also develop a methodology, that relies on structural simultaneous equations, to impute allowance cost pass-through by dispatch technology. We find that pass-through was (more than) complete in seven major European markets for both coal- and gas- fired power generation.

KEY WORDS: Fuel-switching, European Union, Allowances, Emission Trading, Gas, Coal, Electricity

JEL CODES: Q48, C36, C51, C54, L23, H22

*Bai: Zhongnan University of Economics and Law, School of Finance, 182 Nanhu Avenue, Wuhan, 430073 Hubei, China, yiyi.bai@zuel.edu.cn.

†Okullo: Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, P.O Box 601203, D-14412 Potsdam, Germany, okullo@pik-potsdam.de.

1 Introduction

The European Union (EU) Emissions Trading Scheme (ETS) is a core policy tool through which EU lawmakers hope to achieve deep cuts in carbon emissions. Presently, it is the second-largest by regulated emissions, largest by monetary value, and arguably the most advanced carbon emission trading scheme in the world. Since its inception it has, however, been plagued by volatility and for the most part secularly falling allowance prices for which observers have noted: disrupts this carbon market's proper functioning and disincentivizes emission reduction efforts (e.g., Fuss et al., 2018). With the Market Stability Reserve (MSR) now in place, the hope is that the availability of allowances will be regulated as to sustain a strong link between allowance prices and abatement activity. Yet, it still remains unclear to what extent regulated entities in the power industry across Europe responded to the EU ETS prior to the MSR reform. Did the low and volatile prices observed during Phase II and most of Phase III[1] curb generators responses?

The aim of this article is, therefore, to assess whether power generators' responses to the EU ETS aligns with expectations as prescribed by economic theory. This issue is not only of historical significance specific to the EU, but crucially informs the role that ETSs can play in global emission reduction efforts. While prior empirical research has shed some light on this, a clear picture is yet to emerge because: (i) there are contradictory findings on whether allowance prices do in fact reflect marginal abatement costs (e.g. Aatola et al., 2013; Creti et al., 2012; Hintermann, 2010; Koch et al., 2014; Lutz et al., 2013; Mansanetbataller et al., 2007; Rickels et al., 2015), and (ii) the state of allowance cost pass-through in European economies is still relatively unknown as: on the one hand, pass-through tends to be imputed with error when the marginal unit is unobserved (for discussion and some results see: Guo and Gissey, 2021; Fell et al., 2015; Sijm et al., 2006, 2008); on the other hand, consistent estimates where the marginal unit is observed are available for only two countries: Germany (Hintermann, 2016) and Spain (Fabra and Reguant, 2014).

Our starting point is a dynamic structural model of fuel-switching subject to emis-

[1] Phase I ran from 2005 to 2007, Phase II from 2008 to 2012, and Phase III runs from 2013 to 2020.

sion compliance. The model's equilibrium conveys that: (i) the allowance price responds to both the gas and coal price because they are fundamental determinants of (short-run) abatement costs, (ii) changes in the allowance price reflect in part the incentive to bank allowances in order to reduce and smooth compliance costs[2], and (iii) allowance costs are passed through to the wholesale power market in proportion to the carbon intensity of the generation technology, making technology-specific inference of allowance cost pass-through possible.

We test the predictions of our decomposed structural model against data from Phase II and III up until May 2018. This window was marked by volatile and low allowance prices and is, therefore, of significance for gaining insight into the functioning and versatility of the EU ETS. Our empirical results confirm a systematic and robust relationship between the allowance price and both of its pre-eminent fundamental drivers: the gas and coal price. As theory prescribes, we find that the allowance price increases in the gas price and decreases in the coal price. Such a tight link is consistent with an active fuel-switching mechanism aimed at curbing emission costs. Intriguingly, this finding is the exception rather than the norm, as prior literature (see Hintermann et al., 2016, for a review) typically finds that gas (but not coal) prices help explain movements in allowance prices. We provide a resolution to this "coal puzzle" by jointly elucidating the influence of confounding factors and aiming for efficient estimators.

The relationship between the allowance price, and both the gas and coal price is also robust to conditioning for allowance banking incentives. Theory predicts that allowances banked for compliance help generators reduce deterministic and smooth stochastic shocks in compliance costs. The reduction in deterministic compliance costs is the result of banking sustaining inter-temporal arbitrage such that under certainty, the allowance price is capped to rise at rate of discount (Kling and Rubin, 1997). In the presence of uncertainty, banked allowances have additional value because they help generators smooth shocks in compliance costs (Schennach, 2000). Since these incent-

[2]From the perspective of an agent who holds allowances for financial rather than abatement purposes, an alternative interpretation is that changes in the allowance price, besides variation in the investment opportunity set, reflect the expected asset appreciation and the allowance risk premium (Merton, 1973).

ives of allowances banking are ultimately reflected in variations in the allowance price, we empirically condition for their influence via the risk-free rate and a measure for risk. The latter turns out to empirically explain variations in the allowance price, affirming that holding an allowance helps hedge compliance cost shocks. This finding notwithstanding, both the gas and coal price continue to robustly explain the allowance price in the presence of these additional covariates.

Our investigation climaxes with an analysis on pass-through of allowance costs. We empirically recover estimates of average carbon intensity, by either gas- or coal-firing technology, for seven European power markets. To impute allowance cost pass-through, we compare these empirical estimates to corresponding engineering quantities. Our findings indicate that the representative power generator in Germany, the Netherlands, United Kingdom, Belgium, Italy, and Nordpool passes-through emission costs completely. For France, the results suggest more than complete pass-through. Fabra and Reguant (2014) for Spain, and Hintermann (2016) for Germany, report consistent estimates of pass-through for the marginal generator. Like us, they find complete pass-through of allowance costs. We advance this research on pass-through by introducing a novel identification strategy, that applies when the merit order curve is not fully observed or the marginal unit is unknown, such as with futures markets.

Omitted variable and simultaneity bias are two principle threats to identification. The former stems from not observing the technology markup, yet it is plausibly correlated with gas, coal, and allowance prices. And the latter, from the electricity and allowance price potentially being simultaneously determined via an equilibrium mechanism.

Thus, to consistently impute carbon intensities, and hence allowance cost pass-through, we employ instrumental variables. To the extent that generators settle on the degree of coal- (gas-) fired power generation taking as given the degree of gas- (coal-) fired power generation, our theory, via structural simultaneous equations (Wooldridge, 2016, Ch. 16), establishes that the gas (coal) price is an excluded instrument that can help uniquely identify allowance cost pass-through specific to coal (gas) dispatch. This gives us our first of two instruments. The other, is EU ETS regulatory and policy shocks

whose announcements or outcomes were not fully anticipated by the allowance market. Besides enabling us to conduct tests of instrument validity, this additional instrument provides a source of exogenous variation for more precisely identifying allowance cost pass-through. We assume that the "surprise" policy announcements or impacts, affect the electricity market only through their impact on allowance prices.

Our article thus brings together three lines of inquiry with the goal of obtaining a wholistic picture of the effects of EU ETS on the representative power generator. Firstly, the research on the fundamental drivers of the allowance price (e.g. Koch et al., 2016; Rickels et al., 2015; Koch et al., 2014; Aatola et al., 2013; Creti et al., 2012; Hintermann, 2010). Secondly, that on the dynamic properties of allowance prices (e.g., Chevallier and Sévi, 2014; Benz and Trück, 2009; Paolella and Taschini, 2008), and thirdly, the literature on the pass-through of allowance costs (e.g., Guo and Gissey, 2021; Hintermann, 2016; Fell et al., 2015; Fabra and Reguant, 2014). More broadly, this article contributes to the nascent literature that evaluates the causal effects of the EU ETS. Bayer and Aklin (2020) estimate that despite low allowance prices, between 2008 and 2016, the EU ETS reduced emissions by about 1.2 billion tons. Germeshausen (2020) examines efficiency improvements in fossil fuel use, and attributes 1.5–2 percent annual carbon emissions reductions within the German energy sector to the EU ETS. Our research supports these findings, indicating that although allowances prices might have been low and volatile for a long-time, the EU ETS remained (perhaps surprisingly) functional, meaning policies such as the MSR that aim to raise prices can further boost its capacity to curb emissions.

The rest of this article is organized as follows. Section 2 presents the dynamic structural model that guides our empirical evaluation. We derive the equilibrium conditions of this model, discuss how to take it to the data, and describe the empirical challenges that must be overcome for meaningful inference. Section 3 discusses the data collected for the empirical exercise and presents some descriptive statistics. The empirical link between the allowance price and its drivers is presented in Section 4, while 5 evaluates the pass-through of allowance costs. Section 6 concludes.

2 Theory model of power generation under emission compliance

This section introduces our theory model of fuel-switching for emissions reductions and derives the equilibrium that guides the empirical analysis. We consider a representative power generator whose activities span the electricity, allowance, natural gas, and hard coal markets[3]. While the generator can mark up the electricity price—and therefore knows the demand function—we conversely assume that the allowance and fuel input markets are perfectly competitive.

2.1 Model

Because our interest is in emissions abatement through carbon-based fuel switching, we assume that the generators' portfolio consists only of gas- and coal- fired power plants. Moreover, we assume there is sufficient capacity to sustain equilibrium production and consequently abstract from capacity adjustment decisions. This abstraction is not an issue since for the empirical estimation, any capacity rents can be seen as embodied in the unobserved markup.

At any time t, the generator determines its generation of electricity, $e(t)$, by optimally switching between gas-, $g(t)$, and coal-, $c(t)$, firing technology. So as to comply with emissions control requirements, the generator submits enough allowances to the regulator to cover its emissions of the period. Allowances submitted for compliance are either contemporaneously acquired in the market or are recovered from the generator's banked stock, $a(t)$. Moreover, in case the generator has banked allowances in excess of immediate compliance needs, we accommodate the possibility that banked allowances are tradable in the market for profit.

We posit that the current electricity price $p^e(\cdot, t)$, and allowance price, $p^x(t)$, are both known with certainty, but that future prices are uncertain. For electricity prices, this uncertainty can stem from imprecise knowledge regarding future fluctuations in residual electricity demand. In contrast, stochastic fluctuations in the allowance price

[3]Hard coal (unlike lignite or brown coal) is extensively traded internationally, and its prices are internationally quoted.

may stem from regulatory and policy shocks or the activities of speculators. It is also possible that factors driving uncertainty in fossil electricity demand may correlate with those driving uncertainty in allowance prices. Such an extension only slightly modifies the equilibrium conditions, however, and turns out to not have a statistically meaningful impact on the reported results.

Formally, the representative fossil fuel generator seeks to maximize the expected net present value of profit, $J(\cdot)$, as defined in the following optimization program:[4]

$$\max_{g(\cdot),c(\cdot),x(\cdot)} \mathbb{E}_k \int_{t\geq k}^{\infty} \rho(t) \left\{ p^e\left(E(t), y(t)\right) e(t) - p^g(t)g(t) - p^c(t)c(t) - p^x(t)x(t) \right\} dt \quad (1)$$

s.t.

$$\dot{a}(t) = x(t) - \phi^g g(t) - \phi^c c(t) \quad (2)$$

$$dy(t) = \alpha^y\left(y(t), t\right) dt + \sigma^y\left(y(t), t\right) dz^y(t) \quad (3)$$

$$dp^x(t) = \alpha^x\left(p^x(t), t\right) dt + \sigma^x\left(p^x(t), a(t), t\right) dz^x(t) \quad (4)$$

where $\rho(t) = \exp\left(r\left(k - t\right)\right)$ and $e(t) = e^c\left(c(t)\right) + e^g\left(g(t)\right)$. k denotes the initial period, r the risk-free discount rate, ϕ^g the unit emission intensity for the natural gas input, and ϕ^c the unit emission intensity for the coal input. $E(t)$ is demand, which due to market clearing, is identical to supply. $e^g\left(g(t)\right)$ represents power supplied using gas-firing technology and $e^c\left(c(t)\right)$ power supplied using coal-firing technology. $p^e\left(E(t), y(t)\right)$ is the inverse demand function, which besides $E(t)$, also depends on a stochastic exogenous shifter, $y(t)$. $p^g(t)$ and $p^c(t)$ are the gas and coal price, respectively. Since we restrict the generator from borrowing allowances, we have that $a(t) \geq 0$, but permit trades such that $x(t) \gtreqless 0$. The generator knows the market demand curve and thus can markup the power price. But such a generator is also small compared to the size of the fuel input and allowance market and therefore takes these input prices as given (cf. Fabra and Reguant, 2014; Hintermann, 2016; Koch et al., 2016). In equilibrium, firms' expectations about the growth rate of the allowance price ensure that the cumulative

[4]One way to look at this model is to see the power company as split into several arms. One arm does long-term investing to ensure there is more than sufficient capacity for the generation arm. The generation arm, on the other hand, is involved in the day-to-day decisions of supplying power given excess installed capacity. Our model is for the generation arm.

supply of allowances exactly matches the cumulative emissions cap.

Objective (1) says that the generator chooses gas and coal inputs, as well as the transaction of allowances such that the expected discounted sum of net profits is maximized. Equation (2) indicates that when the generator over-complies in the current period, excess allowances are banked for future compliance or trading. Equations (3) and (4) track the stochastic evolution of both the fossil fuel demand shifter and the allowance price. Here, the term $\alpha^y(\cdot)$ ($\alpha^x(\cdot)$) represents the drift and $\sigma^y(\cdot)$ ($\sigma^x(\cdot)$) the volatility of the demand shifter (allowance price). While the drift, $\alpha^x(\cdot)$, is not affected by the stock of banked allowances, volatility is potentially affected.[5] $z^y(t)$ and $z^x(t)$ are Wiener processes associated with the demand shifter and allowance price, and propagate as: $dz^y(t) = \varepsilon^y(t)\sqrt{dt}$ and $dz^x(t) = \varepsilon^x(t)\sqrt{dt}$ where $\varepsilon^y(t)$ and $\varepsilon^x(t)$ are time-wise serially uncorrelated normal random variables with zero mean and unit variance.

2.2 Equilibrium

The above defined model can be solved using stochastic dynamic programming. On the basis of the fundamental equation for optimality, and henceforth suppressing the time index for convenience, the Bellman equation (e.g., Chang, 2004, pg. 113 - 167) is:

$$rJ = \max_{g,c,x} \left\{ \pi + \dot{a} J_a + \alpha^x J_{p^x} + \alpha^y J_y + \frac{1}{2}(\sigma^x)^2 J_{p^x p^x} + \frac{1}{2}(\sigma^y)^2 J_{yy} \right\} \tag{5}$$

where $\pi = p^e(E, y)\, e - p^c \cdot c - p^g \cdot g - p^x \cdot x$. J_a, J_{p^x}, $J_{p^x p^x}$, J_y, and J_{yy}, are partial derivatives of the value function, J, with respect to a, p^x, or y.

Differentiating the Bellman equation with respect to g, c, and x on the assumption that generators take each others choices as given, i.e., assuming strategies that satisfy Cournot-Nash, yields:

$$\pi_g = \phi^g J_a, \qquad \pi_c = \phi^c J_a, \qquad \pi_x = -J_a \tag{6}$$

[5] Although a small generator is unlikely to influence the (expected) price, it could still have a substantial impact on market volatility. For instance, during Phase I of the EU ETS, news that some Dutch firms were over allocated led to substantial volatility in the allowance price without affecting the trend of the price. Prices only began to drift after it become clear that other European firms were over allocated.

where

$$\pi_g = p_E^e\left(E,y\right)e_g^g\left(g\right)e^g\left(g\right) + p^e\left(E,y\right)e_g^g\left(g\right) - p^g \tag{7a}$$

$$\pi_c = p_E^e\left(E,y\right)e_c^c\left(c\right)e^c\left(c\right) + p^e\left(E,y\right)e_c^c\left(c\right) - p^c \tag{7b}$$

$$\pi_x = -p^x \tag{7c}$$

π_g and π_c denote marginal revenue net of fuel cost for gas- and coal- fired power generation respectively, whereas π_x is the marginal cost of an allowance. $p_E^e\left(\cdot\right)$ is the derivative of the inverse demand function with respect to E, whereas $e_g^g\left(g\right)$ and $e_c^c\left(c\right)$ define the marginal product of gas and coal technology. In practice, $e_g^g\left(g\right)$ and $e_c^c\left(c\right)$ are given by the efficiency of gas- and coal- fired power plants.

The first and second sub-equations of (6) imply that the generator burns gas or coal inputs to ensure that net marginal revenue is equal to the emission-intensity adjusted shadow value of the banked stock of allowances. The third sub-equation requires that the generator use up its bank of allowances to the point where the incremental value of having an additional allowance in the bank equates to the allowance price.

Using (7a)-(7c) in (6), the first order necessary conditions can be compactly expressed as:

$$p^e = \frac{1}{e_g^g\left(g\right)}p^g + \frac{\phi^g}{e_g^g\left(g\right)}p^x + \chi^g \tag{8a}$$

$$p^e = \frac{1}{e_c^c\left(c\right)}p^c + \frac{\phi^c}{e_c^c\left(c\right)}p^x + \chi^c \tag{8b}$$

where $\chi^g = -p^e e^g\left(\epsilon E\right)^{-1}$ and $\chi^c = -p^e e^c\left(\epsilon E\right)^{-1}$. χ^g and χ^c are the generator's technology-specific markup and $\epsilon\left(<0\right)$ the elasticity of demand for electricity.[6] Equation (8a) describes the ceteris paribus optimality condition for gas-fired power generation/dispatch, and (8b) the equivalent condition for coal-fired power dispatch. While each equation can stand on its own, in equilibrium with both fuels supplied, the two become linked via the market-clearing electricity and allowance price.

The cheaper source of power that is dispatched first earns the greater markup. The

[6]Note that the elasticity is given by: $\epsilon = \frac{p^e \partial E}{E \partial p^e}$

most costly technology supplying in the merit order earns the least mark-up and is regarded to be the market-clearing technology. Fabra and Reguant (2014) and Hintermann (2016) both estimate allowance cost pass-through for the marginal operator. To achieve this, they build hourly merit order curves using detailed plant-level data. Our empirical strategy in contrast to theirs is agnostic about the marginal unit/technology. For identification, we instead exploit the order and rank condition classical to structural simultaneous equations (e.g. Wooldridge, 2018, Ch. 16.3). Essentially, Equations (8a) and (8b) say that because the generator ensures that the technology being dispatched (more than) breaks-even independently of the competing technology's fuel price, the gas (coal) price is a plausible excluded instrument for imputing allowance cost pass-through specific to coal- (gas-) fired power generation. The rank condition, which is both necessary and sufficient for identification, requires $\phi^g / e_g^g(g) \neq 0$, $\phi^c / e_c^c(c) \neq 0$, and $\phi^g / e_g^g(g) \neq \phi^c / e_c^c(c)$.

Thus, the quantities of interest in the empirical exercise are ϕ^g, ϕ^c, e_g^g, and e_c^c. We assume that these parameters are stable over extended periods of time, something supported by the engineering data (see Table 1).

Western Europe has substantial capacity for coal-gas fuel switching. In all the power markets that we evaluate in the empirical excercise, both gas- and coal- fired power are simultaneously fed into the grid on a minute-by-minute basis clearing at the same p^e. Thus, we can express p^x as a function of p^g and p^c by equating the left hand sides of Equations (8a) and (8b):

$$p^x = \frac{e_g^g(g)\, e_c^c(c)}{\phi^c e_g^g(g) - \phi^g e_c^c(c)} \left[\frac{1}{e_g^g(g)} p^g - \frac{1}{e_c^c(c)} p^c + \chi^g - \chi^c \right] \tag{9}$$

In practice $\phi^c e_g^g(g) > \phi^g e_c^c(c)$ in which case, (9) predicts that the allowance price increases in p^g but decreases in p^c, all other effects constant. The explanation for this is that an increasing gas (coal) price increases the demand for coal (gas), which leads to a higher (lower) demand and hence price for allowances. In the empirically relevant case where gas-fired power generation is more efficient than coal-fired power generation, the equation predicts that the allowance price will be more responsive to a change

in the coal price than to a change in the gas price.

Expectations regarding the growth rate of the allowance price are a key driver of long-run abatement. For a positive stock of banked allowances, the model (1)-(4) yields the following expression for the expected change in the allowance price:[7]

$$(1/dt)\,\mathbb{E}dp^x = rp^x + \sigma^x\sigma_a^x J_{p^x p^x} \tag{10}$$

This equation says that on average, the change in the allowance price can be decomposed into a scarcity and volatility premium.[8] In banking an allowance, its holder expects it to on average return the risk-free rate, r, plus the asset-specific risk premium, $\sigma^x\sigma_a^x J_{p^x p^x}/p^x$. The former is associated with the fact that in a fully deterministic environment, banking allowances helps reduce compliance costs by capping the growth rate of the allowance prices to r (Rubin, 1996). The stochastic component, on the other hand, captures the fact that under uncertainty, banking helps smooth compliance costs by protecting the holder from allowance price shocks (Schennach, 2000).

Note that if firms choose not to explicitly hold allowances, in order to offset non-compliance risk, they will purchase derivative instruments which due to arbitrage are expensed per Equation (10). Conditional on $\sigma_a^x > 0$, the expected rate of change of the allowance price is increasing (decreasing) in price volatility for $J_{p^x p^x} > 0$ ($J_{p^x p^x} < 0$). The possibility of substituting between coal and gas, combined with the fact that allowances can be traded in the market suggests that the case[9] $J_{p^x p^x} > 0$ is empirically more plausible. These two conditions together, i.e. $\sigma_a^x > 0$ and $J_{p^x p^x} > 0$, mean that we can expect the risk premium to be positive.

[7] To derive this expression, we differentiate (5) with respect to a, and use the equilibrium condition (6) to obtain a first expression. Differentiating the value function $J(a, y, p)$ with respect a, and applying Itô's differential generator $(1/dt)\,\mathbb{E}_t d(\cdot)$ yields a second expression. Combining these two expressions, gives this equation.

[8] In an extended model, this expression will include additional terms that also describe changes in the investment opportunity set as in e.g. Ghysels et al. (2005) and Merton (1973).

[9] This essentially means that the value function is convex in the allowance price such that the generator—who is either a buyer or seller of allowances—has the incentive to switch to the less polluting technology at increasingly higher allowance prices.

Table 1: Engineering-based power plant efficiencies, heat rates, and emission intensities for some countries in Europe.

	Germany		Netherlands		U. Kingdom		Belgium		Italy		France		Spain		Portugal	
	2005	2016	2005	2016	2005	2016	2005	2016	2005	2016	2005	2016	2005	2016	2005	2016
PANEL A: power plant efficiency in percent ($e_g^g(g)$ and $e_c^c(c)$)																
Natural gas	0.404	0.444	0.418	0.503	0.466	0.489	0.484	0.512	0.454	0.464	0.480	0.480	0.629	0.556	0.507	0.371
Coal	0.377	0.371	0.418	0.411	0.355	0.35	0.379	0.409	0.372	0.380	0.368	0.383	0.375	0.385	0.327	0.388
PANEL B: heat rate in MWh Fuel per MWh Electricity ($1/e_g^g(g)$ and $1/e_c^c(c)$)																
Natural gas	2.472	2.252	2.390	1.990	2.146	2.045	2.065	1.953	2.203	2.155	2.083	2.083	1.589	1.798	1.973	2.699
Coal	2.655	2.696	2.390	2.433	2.817	2.857	2.636	2.446	2.690	2.632	2.717	2.612	2.669	2.596	3.054	2.576
PANEL C: emission intensities after combustion in tCO2 per MWhe ($\phi^g/e_g^g(g)$ and $\phi^c/e_c^c(c)$)																
Natural gas	0.455	0.415	0.440	0.366	0.395	0.376	0.380	0.359	0.405	0.397	0.383	0.383	0.292	0.331	0.363	0.497
Coal	0.904	0.918	0.814	0.829	0.959	0.973	0.898	0.833	0.916	0.896	0.925	0.890	0.909	0.884	1.040	0.877
Coal/Gas ratio	1.987	2.215	1.850	2.263	2.585	2.429	2.362	2.318	2.259	2.259	2.413	2.320	3.108	2.672	2.864	1.766

Notes: Country-level power plant efficiency data except for France provided by ENERDATA. Country-level power plant efficiencies for France from RTE. Emission intensities before combustion are standard quantities: ϕ^g=0.18404 and ϕ^c=0.34056. The heat-rate is the inverse of the power plant efficiency. Carbon intensities after combustion is the product of emission intensity before combustion and the heat rate. Note that $\phi^c = 0.34056$ pertains to the emission intensity of hard (not brown) coal.

2.3 Testable implications from the decomposed structural model

The primary objective of this paper is to examine whether empirical data on the EU ETS regulated power generating sector is consistent with the (decomposed) structural model above. Three predictions coming out of this model are testable. First, does the allowance price systematically respond to both the gas and coal price as its short-run fundamental drivers? Second, conditional on observing a positive bank, do either or both the risk-free and risk return help explain variations in the allowance price? Third, to what extent do power generators pass-through allowance costs to the wholesale power market?

Equation (9) pertains to the first issue of whether the allowance price responds to both the gas and coal price as its preeminent drivers. A positive (negative) empirical relationship between the allowance price and the gas (coal) price would be qualitatively consistent with the theory. And by plugging the engineering parameters shown in Table 1 into Equation (9), we deduce that the empirically estimated coefficients on the gas and coal price ought, in absolute terms, to be of comparable magnitude. For instance, efficiencies of $e_g^g = 0.5$, $e_c^c = 0.4$, and emission intensities before combustion of $\phi^g = 0.18404$, $\phi^c = 0.34056$, yield a coefficient of 4.14 (5.17) on the gas (coal) price. These magnitudes of course disregard important inefficiency-inducing drivers such as trading costs and thin trading (Jaraitė et al., 2010; Heindl, 2017; Hintermann et al., 2016). Such "frictions" will generally depress these coefficients, but it is not readily quantifiable to what extent. Moreover, as we discuss in Section 2.4, if the gas and coal

price are endogenous, their empirically imputed coefficients are biased towards zero.

Equation (10) pertains to the second issue. In theory, banked allowances on average return the risk-free rate plus the asset-specific risk premium. Thus, all else constant, both the observed risk-free rate and the risk premium, the latter of which can be proxied for using conditional volatility (Ghysels et al., 2005), should in practice help explain variations in observed allowance prices. In the empirical exercise and consistent with the notion of a time-varying investment opportunity set (Ghysels et al., 2005; Merton, 1973), we introduce those two quantities, besides variations in the gas and coal price, as additional explanatory variables for variations in the allowance price. If the risk-free rate helps explain variations in the allowance price, per the theory, it would be indicative of a dynamic compliance value of banking allowances, separate from the value of compliance cost smoothing. If the riskiness term helps explain the allowance price, this would suggest a compliance cost smoothing value of banked allowances.

Using Equations (8a) and (8b), we can evaluate the third issue: to what extent do generators pass-through allowance costs? Foremost, this requires that the coefficient on the allowance price, i.e., carbon intensity after combustion[10], is statistically different than zero such that emission costs "bite." We can subsequently check whether the empirically estimated coefficients for technology-specific carbon intensity, i.e. $\phi^g / e_g^g(g)$ and $\phi^c / e_c^c(c)$, line up with those observed in the engineering data. For a selection of countries, Table 1 presents these engineering-based carbon intensities. If the empirical intensities are consistently estimated, they must line up with the engineering figures. This in turn implies that per megawatt of electricity (MWhe), generators must on average pass-through the cost of an allowance for gas-fired generation at one-third to one-half the magnitude for coal-fired generation.

2.4 Consistent estimation of the decomposed structural model

Estimating the decomposed structural model set out above faces a number of challenges. Firstly, if prices (electricity, gas, coal, and allowance) are non-stationary and there is no cointegrating relationship amongst them, then statistical inference on oth-

[10]For the rest of the article, whenever we use carbon intensity without qualifying it with before/after combustion, we will be referring to after combustion.

erwise untransformed data is unreliable due to spurious relationships. Secondly, in the presence of heteroskedasticity and autocorrelation, ordinary least squares (OLS) is not an efficient estimator. This can lead to falsely accepting a null hypothesis. Thirdly, OLS is biased and inconsistent if fuel and allowances prices are endogenous. Indeed, we do not observe the markups and there is simultaneity if relevant prices are contemporaneously determined.

To deal with non-stationarity, we difference the data. Yet differencing eradicates information on long-run relationships, such that imputed estimates may be different between levels and between increments. Fortunately, if the relationship postulated in levels is not misspecified and in the absence of endogeneity, the underlying relationship can still be consistently estimated using first differences (Wooldridge, 2018, p 381). When studying the relationship between the allowance price and its fundamental drivers, we make use of both the OLS estimator on log-differenced prices and the Feasible Generalized Least Squares (FGLS) estimator on differenced prices. Log-differencing allows us to stabilize the variance and thereby obtain more adequate standard errors. By contrast, FGLS leads to more efficient estimates than OLS in the presence of heteroskedasticity.

We disregard potential endogeneity when examining the response of the EUA price to both the gas and coal price. This assumption is used (e.g. in Aatola et al., 2013; Hintermann, 2010; Koch et al., 2014; Lutz et al., 2013; Mansanetbataller et al., 2007) That being said, Equation (9) shows that if the markups χ^g and χ^c tend to cancel out, such that the correlation of p^g and p^c to the unobserved residual markup is small, bias is moderated. Moreover, because p^c (p^g) is positively (negatively) related to the residual markup, the imputed coefficients will be biased towards zero, i.e., attenuation bias. Thus, unbiased coefficients can only imply a stronger response of p^x to p^g and p^c.

Since the yield on triple-A European government bonds, a standard measure for the risk-free rate, are close to zero and exhibit little temporal variation in our sample, we cannot identify the true underlying response of allowance prices to the risk-free rate. In our principle empirical evaluations we, therefore, exclude the risk-free rate as a control variable and capture its potential impacts through a regression constant. In

the robustness checks, we follow the finance literature and subtract the observed risk-free rate from allowance price returns to construct a new response variable of excess returns. We contrast empirical results on the remaining parameters under these two response variables.

We address endogeneity when imputing pass-through of allowance costs using two instruments. Besides the substitute fuel's price, i.e., the gas (coal) price in the coal- (gas-) fired generation equation, we also use regulatory and policy shocks that plausibly affect the electricity price only through the allowance price. We introduce the supplemental source of exogenous variation to help more precisely identify pass-through. Moreover, with two instruments (rather than one), we have an over-identified system and thereby can test instrument validity. Pass-through of fuel costs is in general not consistently estimated in this article. Therefore, it is of secondary importance.

We structure our empirical analysis as follows. Section 4 focuses on the reduced-form relationship depicted in (9) and (10). Section 5 on estimating the structural relationship in Expressions (8a) and (8b), and thus imputing the pass-through of allowance costs.

3 Data

The primary dataset is standard and extensively used in the literature. It comprises the month-ahead EUA futures price, Title Transfer Facility (TTF) month-ahead gas futures price, the API2 CIF ARA coal settlement price, as well as month-ahead baseload electricity futures prices for: Germany, The Netherlands, United Kingdom, Belgium, Italy, France, and NordPool. This price data, including the EURO STOXX 50 stock market index, water reservoir levels for Norway, and yields on triple-A short-term governments bonds for Germany are obtained from Thomson-Reuters Datastream. The country selections for electricity futures prices are dictated by availability per our subscription to the platform. Although weather is more consequential for the spot market, unusual anomalies could affect futures trading activity. Therefore, we gather corresponding country data for daily mean temperature, sunshine duration, precipitation, and max-

imum wind-speed from The Royal Netherlands Metrological Institute.

We focus on the futures (rather than spot) market because the greater fraction of allowances and electricity is sold on forward markets. Moreover, the brunt of the analysis uses month-ahead (rather than year-ahead) futures as these more closely reflect actual generation decisions. While all data is collected at the daily frequency, for the empirical analysis we aggregate it and take weekly averages thereby ironing out some of the volatility associated with daily data. We take averages (rather than select a single trading day per week) in order to retain most of the information contained within the time series. Does aggregating the data lead to bias? No. Our findings carry through when using daily data, except for the observation that distributed lags are needed to capture the full price impact of coal on allowances prices. The other key advantage of using the lower frequency weekly data, is that we can use the higher frequency daily data for constructing our risk measure per the MIDAS framework as pointed in Section 4.4.

We limit the analysis primarily to Phase II (2008-2012) and III (2013-2020) since banking of Phase I (2005-2007) allowances to future phases was prohibited. This, in combination with an overbanked market led to a secular collapse in EUA prices in the last half of Phase I, developments that are bound to have decoupled allowance prices from fuel-switching fundamentals. To explore the robustness of our primary findings, we also conduct analysis using Zeebrugge and National Balancing Point (rather than TTF) gas futures, report results based on alternative estimation windows, and for the analysis in Section 4 on drivers of the EUA price, also consider futures with year-ahead (rather than month-ahead) maturities. Summary statistics and a description of the price data used is provided in Tables A.2 and A.3.

Fan et al. (2017), Koch et al. (2016), and Conrad et al. (2012) find that regulatory updates and policy announcements concerning the EU ETS have a substantial impact on the allowance price. Accordingly, we control for such impacts when explaining variation in the EUA price. Figure 1 marks some notable historical events while Table A.1 provides the complete list of landmark EU ETS events for which we introduce "event" dummies.

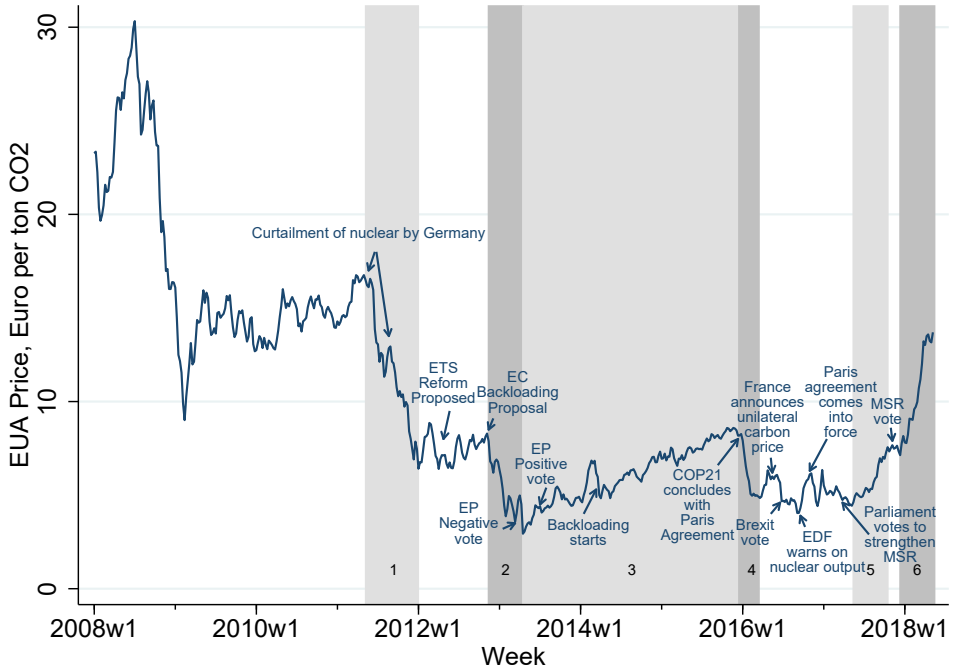


Figure 1: Evolution of the EU ETS allowance price, some historical events, and regimes for notable structural shifts in the trend of the allowance that were spurred by a policy disclosure directly relating to the EU ETS.

Yet, Figure 1 also conveys that persistent new trends emerge in the aftermath of several EU ETS-related policy disclosures. Among other reasons, such trending can result from momentum trading and asymmetry in the capacity for distinct market participants to assimilate or react to new information. Thus, we introduce "trend/duration" dummies to condition for this prolonged structural adjustment. Because it is difficult to judge when the influence of trend-inducing shocks eventually dissipate, we for simplicity assume that when the price reverses/changes trend, the effects of the original policy-induced shock (or any subsequent shocks that did not induce changes in the trend) have dissipated. For shocks originating directly from EU ETS related policy disclosures, this implies six trend dummies indicated by the shaded regions of Figure 1. Including the two trending periods at the outset of the sample window (see Figure A.1), associated with the great recession and European debt crisis, does not change our empirical findings.

We test the presence of a unit root in the EUA, gas, and coal prices (as well as in their log-transformations). The augmented Dickey-Fuller test fails to reject presence of

a unit at the five percent significance level. For electricity prices, a unit root is rejected only for France and Belgium at the five percent level of significance. We also test the presence of a cointegrating relationship using the Johansen test. For the full sample window, there is support for cointegration with the EUA price regressed on natural gas and coal prices but find none over various subsamples inclusive of reducing the full sample by only a single year. To jointly address non-stationarity and heteroskedasticity, we use log-differenced prices in Section 4 when examining the predictors of the EUA price. Notwithstanding, we also report in the same section, results from using FGLS on differenced prices. To impute the pass-through of allowance costs, Section 5 uses differenced prices.

4 Explaining variations in the allowance price

This section examines the empirical response of the allowance price to variations in both the gas and coal price. In addition, it looks at the effects of the of the risk-free rate and risk premium.

4.1 Baseline regressions

Table 2 presents elasticities from regressing the log-differenced allowance price on the log-differenced gas and coal price, and vectors of suitable controls. We use first-differences to ensure the series are stationary. Moreover, log-transforming the prices helps stabilize the variance, leading to more adequate standard errors.[11] Furthermore, with returns (rather than price differences) we can (i) directly compare our results to a majority of prior findings (e.g., Aatola et al., 2013; Lutz et al., 2013) and (ii) summarily evaluate the effects of the risk-free rate and risk-premium within the same specification.

The vectors of suitable controls include calender dummies: week-of-year, month-of-year, and year. These fixed effects control for movements in the allowance price

[11] A Box-Cox specification test fails to reject the log-transformed specification at the 1 percent level, confirming possible improvements as the transformation makes the residuals more closely normal and less heteroskedastic.

Table 2: Elasticity of EUA month-ahead futures to natural gas and coal month-ahead futures.

	Short-run propensities								Long-run propensities	
Covariate	(1) Static	(2) Dynamic	(3) Robust	(4) Non par.	(5) Static	(6) Dynamic	(7) Robust	(8) Non par.	(9)	(10)
Natural gas	0.184***	0.165***	0.216***	0.144***					0.184***	
	(0.056)	(0.052)	(0.049)	(0.043)					(0.058)	
Coal	-0.209***	-0.190***	-0.199***	-0.158***					-0.211***	
	(0.078)	(0.068)	(0.067)	(0.061)					(0.077)	
Relative coal price					-0.191***	-0.172***	-0.212***	-0.150***		-0.191***
					(0.053)	(0.048)	(0.045)	(0.045)		(0.054)
Auto-regressive lag	No	Yes	No	No	No	Yes	No	No	No	No
N	539	539	539	539	539	539	539	539	539	539
adj. R-sq	0.487	0.496	0.478	0.498	0.488	0.497	0.476	0.540	0.570	0.570

Notes: Coefficients in (1)-(8) are impact propensities whereas those in specifications (9) and (10) are long-run propensities (Wooldridge, 2018, p. 607) calculated using specifications (2) and (6). Standard errors for (1) and (5) are New-West heteroskedasticity and serial autocorrelation consistent standard errors. Those in (2) and (6), where the serial correlation is modelled explicitly, are White's heteroskedasticity robust standard errors. One lag of EUA returns proves sufficient in modelling the serial autocorrelation present in EUA price returns. After modelling the serial correlation, we find that the null that errors are homoskedastic cannot be rejected; nonetheless, for conservativeness we report heteroskedasticity robust standard errors. Specification (3) and (7) eliminate outliers based on Cook's distance, and Huber and biweight iterations. Specifications (4) and (8) are based on a non-parametric kernel regression which is robust to functional form misspecification. All regressions are based on the 3-Jan-2008 to 9-May-2019 sample. Week, month, and year dummies are used as controls in all regressions. Dummies are also introduced to control for periods of systematic trending in the allowance price. All specifications except (3) and (7) include dummies to control for week over week returns in excess of 10%. Such drastic movements in the allowance occur only occasionally, and yet can have substantial influence on estimated standard errors. These significant movements can one-for-one be linked to regulatory announcements.
***Significant at 1 percent.

that stem from seasonality in weather, periodic emissions reporting and verification, periodic information disclosures, pattern trading, and phase-specific drivers. We also make use of shock dummies of events listed in Table A.1 and shock duration dummies illustrated in Figure 1. Since OLS is biased and inefficient in the presence outliers, we handle this using two methodologies. Reported in Table 2 columns (3) and (7) are results based on robust regressions designed to statistically handle outliers (Rousseeuw and Leroy, 2005; Verardi and Croux, 2009). This procedure acts as a check against our more flexible approach that uses impulse indicators to condition for uncharacteristic jumps that remain unexplained by our observed controls. In our flexible approach, we consider uncharacteristic price jumps as those in excess of 10% week-on-week—see Figure A.2—but results are unchanged with a tighter 7% or wider 15% margin. Later in Section 4.2 we show that excluding controls for these impulses affects primarily the efficiency of the coal estimate. For now, we exclude the stock market index, Nordic water reserve levels, weather variables, observed risk-free rate, and risk-premium as possible controls. We come back to these in Section 4.2 and 4.4 as part of the robustness checks.

We evaluate two principal specifications relating to the efficiency of gas- vis-à-vis

coal- fired power generation. Let β_c (β_g) be the coefficient on coal (gas) price returns. We consider the possibility where β_g is not necessarily identical to $-|\beta_c|$, and $\beta_g = -|\beta_c|$. From Equation (9), the latter implies gas-fired generation is strictly more efficient than coal-fired generation.[12] Specifications (1)-(4) in Table 2 are based on the unrestricted specification, whereas (5)-(8) are based on the restricted case. Observe that for specifications (1)-(4), the coefficient on the gas and coal price returns are more or less equal in absolute terms. Moreover, the point estimates obtained using the unrestricted specification, match up in absolute terms with those from the restricted model. Not only are these results in line with engineering data that gas-fired power generation is more efficient than coal-fired power generation, but the theory-consistent and statistically significant relationship also implies both gas and coal help explain variations in the allowance price.

The tabulated results indicate that a 1% change in the gas (coal) price on average predicts a 0.18% (0.20%) increase (decrease) in the EUA price. Specification (1) and (5) assume that the relationship between EUA, gas, and coal prices is purely static. Specification (2) and (6) accommodate for dynamic impacts modelled through a single geometric lag. Specifications (3) and (7) exclude dummies for outlying jumps in the allowance price, and instead apply statistical methods to identify these gross outliers and weight them accordingly. Specification (4) and (8) are local linear non-parametric regression estimates that are robust to functional form misspecification (Li and Racine, 2007). And finally, (9) and (10) present long-run propensities calculated from specifications (2) and (6), respectively.

The specifications convey the following. Firstly, that the static representation satisfactorily captures the response of the allowance price to gas and coal prices. Secondly, outlying changes in the EUA price that are explicitly conditioned-out in all specifications but (3) and (7), can in fact be regarded as gross outliers. Thirdly, by comparing

[12]While the partial of that equation yields marginal effects, conversion of partial effects to elasticities and the reverse is straightforward. To see this, assume a stable relationship in levels. It follows that that the partial/marginal effect can be related to elasticities as follows: $\frac{\partial p^x}{\partial p^g} = \beta_g \frac{p^x}{p^g}$ and $\frac{\partial p^x}{\partial p^c} = \beta_c \frac{p^x}{p^c}$ where β_g and β_c are elasticities, and $\frac{\partial p^x}{\partial p^g}$ and $\frac{\partial p^x}{\partial p^c}$ are partial/marginal effects. Now suppose that $\beta_g/\beta_c = -1$ and since in general, $p^g > p^c$, then $\frac{\partial p^x}{\partial p^c} > \frac{\partial p^x}{\partial p^g}$, which implies from Equation (9) that gas-fired power generation is strictly more efficient than coal-fired power generation.

Table 3: Elasticity of EUA prices to natural gas and coal futures' prices.

Variable	month-ahead futures								year-ahead futures		
	(1) Outliers	(2) Dummies	(3) Controls	(4) Zeebrugge	(5) NBP	(6) Phase II	(7) Phase III	(8) Phase I - III	(9) 2008-2018	(10) 2009-2015	(11) 2006-2010
Natural gas (EUR/MWh)	0.229***	0.251***	0.184***	0.132**	0.121**	0.180**	0.203**	0.105***	0.343***	0.422***	0.357***
	(0.081)	(0.077)	(0.057)	(0.053)	(0.048)	(0.082)	(0.086)	(0.039)	(0.077)	(0.095)	(0.099)
Coal (EUR/MWh)	-0.221*	-0.154	-0.211***	-0.188**	-0.179**	-0.219**	-0.220*	-0.151**	-0.203**	-0.375***	-0.236*
	(0.120)	(0.121)	(0.081)	(0.079)	(0.078)	(0.110)	(0.124)	(0.069)	(0.099)	(0.135)	(0.129)
N	539	539	539	539	539	260	279	689	537	364	210
adj. R-sq	0.112	0.035	0.488	0.481	0.481	0.441	0.526	0.508	0.497	0.485	0.416

Notes: The base specification is the static linear regression of Column (1) Table 2.
(1) Baseline model, excluding dummies for outlying changes in the allowance price. (2) Builds on (1) by additionally excluding dummies for periods of sustained trending in the allowance price. (3) Baseline model with EURO STOXX 200 and Norwegian hydro reservoir levels as additional controls, (4) Baseline model with Zeebrugge (rather than TTF) month-ahead natural gas futures, (5) Baseline model with National Balancing Point (rather than TTF) month-ahead natural gas futures, (6) Baseline model for Phase II (3-Jan-2008 to 29-Dec-2012) observations, (7) Baseline with Phase III (3-Jan-2013 to 09-May-2018) observations, (8) Baseline with Phase I-III (3-Jan-2005 to 09-May-2018) observations, (9) Baseline with year-ahead (rather than month-ahead) EUA, Zeebrugge gas, and Inter-continental Exchange (ICE) coal futures for the period 3-Jan-2008 to 09-May-2018, (10) Baseline with year-ahead (rather than month-ahead) EUA, Zeebrugge gas, and ICE coal futures for the period 3-Jan-2009 to 19-Dec-2015, and finally (11) Baseline with year-ahead (rather than month-ahead) EUA, Zeebrugge gas, and ICE coal futures for the period 27-Jun-2006 to 29-Dec-2010. All standard errors are robust for arbitrary forms of serial correlation and heteroskedasticity.
***Significant at 1 percent; **Significant at 5 percent; *Significant at 10 percent.

(4) and (8) to the associated parametric specifications, that no meaningful non-linear response exists that warrants (further) explicit modelling. We confirm this also using Ramsey's Regression Specification Error Test and Tukey's LINK test.

Our finding that the variations in the EUA price are related to variations in both the gas and coal price, is the exception rather than the norm. A great number of studies find that allowance prices consistently respond to variations in the gas price, but not to variations in the coal price (cf., Rickels et al., 2015; Koch, 2014; Hintermann, 2010; Mansanetbataller et al., 2007) and some in fact recover a statistically significant positive (rather than negative) coefficient on the coal price (e.g., Lutz et al., 2013). To our knowledge, only Aatola et al. (2013) recover a negative statistically significant coefficient on the coal price using data observed from early 2005 to the end of 2010. However, the authors sidestep providing a plausible explanation for why their findings diverge from the rest of the literature[13]. In Section 4.3, we detail why we are able to recover a statistically significant and theoretically consistent sign on coal but for now proceed to evaluate the robustness of the above findings.

4.2 Robustness analysis of the statistical importance of coal

Table 3 gives results for how sensitive our baseline findings are to: (1) failing to control for influential outliers within EUA futures returns, (2) additionally not controlling for policy-induced sustained trending in EUA returns, (3) having the EURO STOXX 200 index and Norwegian hydro reservoir levels as controls, (4) using Zeebrugge or (5) National Balancing Point (rather than TTF) as natural gas futures, (6) separately using Phase II or (7) Phase III sample data for the estimations, (8) using Phase I-III observations, and using year ahead (rather than month-ahead) futures for the sample windows: (9) 3-Jan-2008 to 9-May-2018, (10) 3-Jan-2009 to 29-Dec-2015, and (11) 27-Jun-2006 to 29-Dec-2010.[14] To be conservative, we report only findings for the static linear regression, and for brevity, focus on the unrestricted specification.

The results show that the relationship between the EUA price and the gas and coal price is robust. Comparing column (1) in Table 3 to the baseline results in column (1) of Table 2, it is apparent that the significance level for coal is weaker in the former. In particular, although coal is (still) statistically significant at $p < 0.1$, the loss of significance from one percent suggests that uncharacteristic outlying changes, few as they may be, have undue influence on being able to isolate the statistical association between the allowance and coal price. We also see that the adjusted R-squared drastically drops, confirming that a better model is one that conditions for these uncharacteristic changes. Column (2) extends column (1) by additionally excluding trend dummies. Here coal is only significant at the 20 percent level. An intermediate regression that excludes the trend (but not the outlier) dummies yields points estimates 0.189 (-0.146) significant at $p < 0.01$ ($p < 0.1$) for gas (coal). These findings demonstrate that failing to jointly condition for outliers and structural trends is especially detrimental for identifying the statistical importance of coal explaining the allowance price.

The inclusion of the EURO STOXX 200, Norwegian hydro reservoir levels, and weather variables as controls to the baseline specification, does not have a meaningful impact on the estimates. Indeed, we find these controls to be statistically insignificant,

[13] More specifically, they report that their sample size and data is different and yet studies that used data that (to varying degrees) overlaps their window of estimation find different results.

[14] The regression in Column (11) rans from 27-Jun-2006 because this is as far back out as our year-ahead coal futures price data stretches to.

contrary to what has been reported elsewhere in the literature (see e.g. Aatola et al., 2013; Hintermann, 2010; Lutz et al., 2013; Mansanetbataller et al., 2007; Rickels et al., 2015). Differently than these studies, we introduce controls for seasonality as well as for regulatory and policy shocks. Not only are these controls plausibly exogenous, but they condition-out shifts in allowance demand that would otherwise be missed. It is on this account that our results suggest that stock prices are a poor proxy for economic activity, and that Norwegian hydro reservoir levels as well as weather variables do not on average explain movements in the allowance price.

From columns (4) and (5), it is clear that the response of EUA prices is robust across the natural gas futures trading platform. More specifically, the results that use Zeebrugge or NBP data are broadly consistent with those that use TTF data. We also see from columns (6)-(8) that the relationship persists over different estimation windows, with Phase II (column (6)) and Phase III (column (7)) estimates being more or less similar to one another. Phase I-III point estimates (column (8)) are quantitatively smaller than those of the other phases, however.[15] Still, this difference does not change the qualitative and statistical finding that there exists a robust theoretically consistent relationship between the EUA price and both the gas and coal price. The results that use year-ahead futures as reported in columns (9)-(11) are also by and large consistent with those that use month-ahead data.

Unlike most other authors, Hintermann (2010) and Koch et al. (2014) use differenced (rather than log-differenced) prices. Both authors fail to reject the null that the coal price has no impact on EUA prices. Does our data lead to similar conclusions? The OLS estimator applied to price-differenced data indicates that a €1 change in the gas (coal) price is associated with an increase (decrease) of €0.11 (€0.15) in the EUA price. Using heteroskedasticity and autocorrelation consistent standard errors, gas is significant at $p < 0.01$ and coal is significant only at the twenty percent level. The FGLS estimator is asymptotically more efficient than the OLS estimator in the presence

[15] This may in part be explained by the way in which the allowance price series is constructed. Because Phase I allowance prices collapsed to zero for nearly half of the phase, we construct a modified "month-ahead" allowance futures price by concatenating month-ahead futures prices that run to the end of 2005, with December 2008 futures to cover the period 2006-2008, and then month-ahead futures once again from 2009 onward.

heteroskedasticity (Wooldridge, 2018, p. 412, p. 480).[16] This estimator indicates that coal is statistically significant at the five percent significance level. Thus, our identified responses are robust to using log-differenced or differenced futures prices.

4.3 Why has coal's statistical impact been so elusive?

Our findings so far establish that the response of the EUA price to variations in both the gas and coal price is doggedly robust. Given this finding, the natural question to ask is: why has the economic and statistical importance of coal been so elusive in the extant literature?

Column (1) Table 3 shows how failing to control for uncharacteristic outlying changes in the allowance price weakens the statistical significance of coal. Secondly, Column (2) confirms that additionally excluding controls for news-shock-induced trending in the allowance price is also detrimental for efficiency of the coal estimate. A restricted estimation where $\beta_g = -|\beta_c|$ but excludes both the outlier and trend dummies yields a point estimate of 0.223 significant at $p < 0.01$. That coal's statistical significance is more fragile in the unrestricted specification points towards a nuanced problem that relates to the quality of coal relative to natural gas data. We delve into this issue next.

Wooldridge (2018, p 88) discusses how the statistical properties of a regressor affect the precision of its point estimate. When residuals are homoscedastic, the precision of a regressor's point estimate is determined by (i) its variability as gauged from the regressor's variance, and (ii) the unique variation that it exhibits relative to other contemporaneous regressors[17]. Little variation in both dimensions will lead to imprecision in the estimated sampling distribution of the regressor's point estimate. Moreover, with time series, persistence (or high autoregression) is an important dimension that readily becomes problematic in the absence of homoscedasticity. More specifically, for a regressor with high persistence relative to other regressors, the standard deviation of its point estimate is penalized more heavily when actually modelling or correct-

[16] Applying the Breusch-Godfrey test to the corresponding OLS estimation rejects the presence of serial correlation at $p < 0.01$, thus FGLS is consistent under the relatively weak conditions of contemporaneous exogeneity.

[17] Note that (ii) is measured by the R-squared from the regression of the regressor in question (now turned regressand) on the other regressors

Table 4: Variation in natural gas and coal returns

	2008 to 2018						2005 to 2010					
Variable	Obs.	Mean	SD	R. t. Var	AC(1)	AC(2-8)	Obs.	Mean	SD	R. t. Var	AC(1)	AC(2-8)
Coal												
API2	539	-0.04	3.04	1.000	0.330	0.062	311	0.16	3.40	1.000	0.385	0.090
ICE Rotterdam year-ahead	538	-0.02	3.11	1.021	0.265	0.030	210	0.27	3.69	1.088	0.333	0.060
Gas												
ICE TTF month-ahead	539	-0.03	4.12	1.352	0.193	0.025	306	0.19	6.53	1.920	0.235	0.023
ICE TTF year-ahead	539	-0.04	3.27	1.075	0.123	0.012	306	0.16	3.72	1.096	0.347	0.014
ICE Zebrugge month-ahead	539	-0.03	4.49	1.475	0.167	0.011	172	0.21	6.39	1.877	0.196	0.012
ICE Zebrugge year-ahead	539	-0.03	3.20	1.051	0.140	0.024	311	0.17	4.33	1.275	0.116	0.039
ICE NBP month-ahead	539	-0.03	4.75	1.560	0.158	0.013	163	-0.02	6.54	1.920	0.179	0.004
ICE NBP year-ahead	539	-0.03	3.58	1.175	0.231	0.005	163	-0.02	4.53	1.330	0.303	0.031

Notes: Obs.(=Observations), SD(=Standard Deviation), R. t. Var (=Relative total volatility), AC(=Auto Correlation). Relative total volatility is calculated in two parts. First we take the norm of returns and sum over the sample to obtain total variation. In the second step, we divide this 'total variation' by the total variation of the API2 coal price to create the relative part. AC(1) is the autocorrelation at the first lag, and AC(2) is the average autocorrelation for lags (2)-(8). The data is observed as follows. Coal API2 (03-Jan-2005 to 09-May-2018), Coal Continuous Rotterdam (12-Dec-2006 to 04-May-2018), Gas TTF month ahead and year ahead (09-Feb-2005 to 09-May-2018), Gas Zebrugge month ahead (03-Sep-2007 to 09-May-2018), Gas Zebrugge year ahead (03-Jan-2005 to 09-May-2018), Gas National Balancing Point month ahead and year ahead (09-Nov-2007 to 09-May-2018).

ing standard errors for arbitrary forms of serial correlation and (or) heteroskedasticity. This means that without carefully modelling the data generating process of the regress- and, one is more likely to commit a type II error, i.e., failing to reject a null that should actually be rejected.

Table 4 shows that by many measures of variation, coal in our sample has much less variability than natural gas. We see that both the API2 coal price and the year-ahead coal futures price have lower standard deviations than natural gas futures for the two sample periods 2008 to 2018 and 2005 to 2010. We also calculate the relative total volatility. Coal has smaller total volatility than natural gas prices. The table also shows that coal exhibits higher autoregression than natural gas both at the first lag and over the average of the second to eighth lag. With much less variability in coal prices than natural gas prices, it is not surprising to see why coal's statistical significance has been so elusive. Little variation—especially relative to other regressor's—implies that the variance for a regressor's point estimate will tend to be large and may straddle zero, leading to a failure to reject the null even when it should be rejected (Kennedy, 2005). This problem can be overcome by gathering more data or controls. Or as suggested by Kennedy (2005), working with ratios, which here is equivalent to the restriction $\beta_c = -|\beta_g|$.[18] We have used both approaches and confirmed the presence of a theory-

[18]This plausibly is the reason why analysis that use the fuel-switching price—a combination of the gas and coal price in fixed proportions—have tended to find more theory consistent results than those that use the gas and coal price in unrestricted fashion (see e.g., Creti et al., 2012).

consistent and statistically robust response of EUA prices to variations in both the gas and coal price.

4.4 Incentives for banking

Generators bank allowances to reduce and smooth compliance costs. Equation (10) shows that these incentives can be decomposed from allowance price returns.[19] Thus, whether or not we condition for the effects of the risk-free rate and asset-specific risk premium has the potential to impact the results.

Table 5: The responsiveness of EUA returns to the scarcity and volatility premium.

	Phase II-III			Phase II			Phase III		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
MiDaS									
Discount rate		-4.327	0.283		-1.466	0.191		1.231	0.353
		(3.490)	(3.451)		(1.593)	(1.687)		(4.165)	(4.087)
Volatility	87.540***	87.558***	85.093***	51.704	61.448	65.798	90.219***	88.882***	85.814***
	(20.921)	(20.878)	(21.613)	(82.762)	(83.654)	(86.097)	(20.467)	(20.484)	(21.022)
N	539	539	539	260	260	260	279	279	279
adj. R-sq	0.542	0.511	0.497	0.488	0.455	0.431	0.564	0.557	0.550
BIC	3397.016	3395.983	3400.638	1709.021	1707.663	1710.400	1880.212	1879.705	1875.673

Notes: Volatility is predicted using daily squared returns (Ghysels et al., 2006). Column (1) uses the allowance price returns net of the observed risk-free rate as the predictor variable. Column (2) and (3) both use gross returns as the predicted variable. Because column (1) uses excess returns, the constant in this equation is not given the interpretation of a risk-free discount rate and is therefore not reported. Relative to Column (2), Column (3) excludes gas and coal price returns as predictors of returns. Volatility is computed using the step-wise MiDaS, where the duration of a step is the number of trading days in the week—five in the absence of a holiday. Two steps are used for predicting conditional volatility. In all regressions, we control for serial correlation using the lagged returns (cf., Ghysels et al., 2005; French et al., 1987). Standard errors are robust for heteroskedasticity. Because the estimating equation (10) is an equilibrium relationship, we make use of 'year' and 'week of year' as additional controls. Moreover, dummies for uncharacteristic changes in EUA prices, and dummies for periods of systematic trending are also used to control for events or outcomes that were unanticipated by the generator at the time the purchasing decisions was taken.
***Significant at 1 percent.

The constant when regressing allowance price returns on gas and coal price returns, can be seen as the imputed risk-free rate. In keeping with the observed average yield on the German government bond during our sample window, this component is not statistically different than zero (see Table 5). In an alternate treatment, we subtract the observed risk-free rate (i.e. yield on the 3-month German government bond) from allowance price returns. Regressing this excess return on several controls described above that condition for confounding factors while excluding the constant, leaves the coefficient on gas and coal prices practically unchanged. For our sample, the risk-free

[19]Even for an actor who holds allowances for the sole purpose of arbitrage trading, that equation applies since they have to be indifferent between holding allowances or an alternative asset of identical risk characteristic.

rate thus, does not appear to be informative in explaining movements in the allowance price.

To examine the role of the asset-specific risk premium in explaining EUA price returns, we regress gross or excess EUA price returns on conditional volatility (a standard proxy for risk). We construct conditional volatility per the MiDaS framework which exploits the volatility signal embedded within higher frequency observations to predict volatility at lower frequencies (Ghysels et al., 2005, 2006). Here, we predict volatility at the weekly frequency using both daily squared and absolute returns as detailed in Ghysels et al. (2006).

Table 5 presents results for the linear two-step MiDaS with volatility predicted using daily squared returns. Column (1) uses excess returns as the response variable, whereas Columns (2) and (3) use gross returns. Column (1) and (2) include gas and coal price returns as covariates but column (3) excludes them. In (1) and (2), the coefficients on the gas and coal price are practically unchanged from those in Table 2 and therefore are not presented again. The table shows that EUA price returns are increasing in volatility. This association is statistically significant over the whole sample and in Phase III, but not during Phase II. In the context of our theory, the presence of a positive risk-premium means that banking allowances can have value for smoothing adverse allowance price shocks, and thereby help in reducing compliance costs.

As a robustness check, in Appendix E we predict volatility using daily absolute (rather than daily squared returns). Those results support and are inline with the above findings.

5 Do generators optimally pass-through costs of emissions?

A systematic relationship between the allowance price and both the gas and coal price, is consistent with fuel-switching to curb emission costs. Moreover, we have documented that allowance price returns feature a positive risk-premium, implying that banked allowances have value in helping generators smooth and thereby partially

reduce compliance costs. Inline with the third pillar of our decomposed structural model, this section tests the extent to which the representative power generator passes-through allowance costs to the wholesale power market.

The principle estimating equations derive from the structural simultaneous Equations (8a) and (8b). As earlier noted, since we do not observe the the mark-up, it is encapsulated in the error term, meaning OLS yields biased and inconsistent estimates when the error component is correlated with EUA, gas, and coal prices. Estimates are also biased and inconsistent if the EUA price depends on shocks to electricity prices. In these circumstances, consistent estimates can be recovered via three-stage least squares, two-stage least squares, or Instrumental Variables via Generalized Method of Moments (IV-GMM). Obtaining heteroscedasticity and autocorrelation consistent standard errors is more straightforward with the later two. Below we report the IV-GMM results.

We regress in first differences, the electricity price on the EUA price and either the gas or coal price. We use differences since as earlier noted, we fail to reject the presence of a unit root in several of the electricity prices at $p \leq 0.05$ and fail to reject for all at $p \leq 0.01$. We include dummies for week-of-year, month-of-year, and year to control for seasonality and fixed effects. Norwegian hydro reservoir levels, the EURO STOXX 50 index, Brent crude oil prices, and country weather anomalies (rainfall, wind, sunshine, and humidity) are additional controls (cf. Fell et al., 2015).

The two aforementioned instruments we exploit to identify allowance cost pass-through are: (i) EU ETS policy announcements discussed in Section 3 and listed in Table A.1, and (ii) the competing technology's fuel price. To the extent that power generators do not anticipate several of those announcements or their outcomes, the concomitant shocks are arguably independent of the unobserved mark-up. Moreover, since those policies pertain directly to the EU ETS, they plausibly affect the power market only via their impacts on the EU ETS price.

Policy shocks alone cannot identify pass-through, however. For this we need an explanatory variable that enters Equation (8a) but not (8b); and similarly, (8b) but not (8a) (Wooldridge, 2018, Ch. 16). That is, the variable should shift gas- (coal-) fired

power generation if coal- (gas-) fired generation were to be held fixed. Our theory points to the gas (coal) price as a viable instrument for the allowance price in the coal- (gas-) fired power dispatch condition. In practice, generators commit to supplying a fixed amount of power for a pre-established price. The generators need to ensure that the technology being dispatched (more than) breaks-even independently of the competing technology's fuel price. For this reason, the competing technology's fuel price is a compelling instrument.

Parallel to the preceding theory and narrative motivations for our instruments, below we report both the Hansen-Sargan test and the Montiel Olea and Pflueger (2013) F-statistic. These statistical tests are supportive of the validity and relevance of our instruments.

When the marginal unit/technology is unknown, failing to exploit the exclusion restrictions implied in Equation (8a) and (8b), by projecting the electricity price jointly on the EUA, gas, and coal price, fails to identify pass-through. Here, the identification challenge stems from not knowing the extent to which the identified carbon intensity refers to gas- and/or coal- technology.[20] Table A.6 presents estimates for such an exercise. The coefficient on the EUA price is generally close to the emission intensity for gas-fired power generation. This does not, however, mean that pass-through for gas- (coal-) based power generation is complete (less than complete). Our identification strategy resolves this issue.

Table 6 reports both OLS and IV-GMM estimates for the response of the month-ahead electricity futures price to variations in the month-ahead gas, coal, and allowance futures price. Standard errors are corrected for heteroskedasticity and arbitrary forms of serial correlation. The table also reports the p-value for validity of over-identifying restrictions (J test), and the effective F statistic (F_{Eff}) for the relevance of excluded instruments. In all markets, the J test fails to reject the null of valid over-identifying restrictions. Moreover, the F_{Eff} statistic sits well above the consensus threshold

[20] Fell et al. (2015) and Hintermann (2016) attempt to resolve this issue by exploiting additional restrictions on the frequency with which gas or coal is marginal. Their imputed shares differ markedly from those observed in the data. Guo and Gissey (2021) imputes the frequency with which a technology is marginal via its responsiveness to variation in demand. They cannot, however, observe the carbon intensity of imports.

Table 6: Responsiveness of baseload month-ahead electricity futures prices to month-ahead futures gas price and the API2 coal settlement price.

	Electricity price (€/MWh)													
	German		Dutch		United Kingdom		Belgian		Italy		France		Nordpool	
	(1) OLS	(2) IV-GMM	(3) OLS	(4) IV-GMM	(5) OLS	(6) IV-GMM	(7) OLS	(8) IV-GMM	(9) OLS	(10) IV-GMM	(11) OLS	(12) IV-GMM	(13) OLS	(14) IV-GMM
PANEL A: natural gas technology (TTF)														
Allowance (€/tCO2)	0.660***	0.457***	0.608***	0.386***	0.830***	0.696***	0.201	0.484**	0.582***	0.366**	0.654***	0.800***	0.621***	0.360***
	(0.168)	(0.080)	(0.165)	(0.137)	(0.208)	(0.114)	(0.251)	(0.226)	(0.182)	(0.149)	(0.247)	(0.144)	(0.171)	(0.099)
Natural gas (€/MWh)	1.174***	1.164***	1.514***	1.627***	1.309***	1.216***	1.607***	1.577***	1.111***	1.023***	2.026***	2.180***	1.166***	1.193***
	(0.128)	(0.069)	(0.116)	(0.071)	(0.188)	(0.087)	(0.197)	(0.115)	(0.149)	(0.108)	(0.258)	(0.085)	(0.127)	(0.072)
N	539	539	539	539	539	539	505	505	495	495	539	539	537	537
Adj. R-squared	0.420	0.182	0.492	0.309	0.386	0.202	0.363	0.033	0.456	0.112	0.428	0.114	0.417	0.181
J (pvalue)		0.781		0.240		0.313		0.188		0.274		0.240		0.796
F_{Eff}		64.610		64.610		64.610		56.403		51.587		64.610		65.720
PANEL B: natural gas technology (NBP)														
Allowance(€/tCO2)	0.760***	0.582***	0.715***	0.339**	0.931***	0.613***	0.231	0.089	0.628***	0.461***	0.817***	0.721***	0.718***	0.392***
	(0.190)	(0.077)	(0.188)	(0.153)	(0.220)	(0.120)	(0.256)	(0.269)	(0.196)	(0.135)	(0.267)	(0.165)	(0.192)	(0.084)
Natural gas (€/MWh)	0.909***	0.891***	1.263***	1.341***	1.055***	1.033***	1.268***	1.294***	0.883***	0.786***	1.602***	1.811***	0.919***	0.949***
	(0.113)	(0.056)	(0.103)	(0.052)	(0.156)	(0.075)	(0.175)	(0.118)	(0.132)	(0.091)	(0.243)	(0.121)	(0.107)	(0.055)
N	539	539	539	539	539	539	505	505	495	495	539	539	537	537
Adj R-squared	0.392	0.144	0.486	0.296	0.369	0.176	0.354	0.021	0.430	0.072	0.410	0.086	0.394	0.145
J (pvalue)		0.691		0.260		0.160		0.277		0.230		0.193		0.781
F_{Eff}		60.471		60.471		60.471		53.290		51.443		60.471		61.324
PANEL C: natural gas technology (ZEE)														
Allowance (€/tCO2)	0.770***	0.617***	0.742***	0.524***	0.941***	0.826***	0.252	0.421*	0.653***	0.416***	0.848***	0.970***	0.734***	0.531***
	(0.185)	(0.092)	(0.183)	(0.150)	(0.212)	(0.120)	(0.257)	(0.253)	(0.191)	(0.112)	(0.259)	(0.159)	(0.189)	(0.093)
Natural gas (€/MWh)	1.057***	1.062***	1.398***	1.415***	1.234***	1.035***	1.360***	1.248***	0.911***	0.878***	1.792***	1.794***	1.047***	1.039***
	(0.109)	(0.067)	(0.102)	(0.054)	(0.163)	(0.086)	(0.180)	(0.103)	(0.137)	(0.092)	(0.258)	(0.110)	(0.113)	(0.069)
N	539	539	539	539	539	539	505	505	495	495	539	539	537	537
Adj R-squared	0.417	0.179	0.502	0.324	0.402	0.217	0.354	0.019	0.435	0.080	0.420	0.105	0.414	0.178
J (pvalue)		0.684		0.140		0.339		0.216		0.305		0.116		0.652
F_{Eff}		62.757		62.757		62.757		55.714		52.021		62.757		63.679
PANEL D: coal technology (CIF ARA API2)														
Allowance (€/tCO2)	0.913***	0.968***	0.945***	1.282***	1.104***	1.181***	0.507*	1.226***	0.882***	1.125***	1.108***	1.493***	0.871***	0.884***
	(0.192)	(0.117)	(0.192)	(0.162)	(0.197)	(0.148)	(0.277)	(0.244)	(0.199)	(0.197)	(0.286)	(0.170)	(0.201)	(0.123)
Coal (€/MWh)	1.997***	2.031***	2.214***	2.552***	2.491***	2.265***	2.888***	2.915***	2.412***	2.232***	2.881***	3.153***	1.982***	2.036***
	(0.704)	(0.387)	(0.670)	(0.389)	(0.711)	(0.383)	(0.641)	(0.438)	(0.420)	(0.348)	(1.034)	(0.544)	(0.636)	(0.288)
N	539	539	539	539	539	539	505	505	495	495	539	539	537	537
Adj R-squared	0.308	0.028	0.287	0.027	0.281	0.068	0.278	-0.110	0.379	-0.012	0.319	-0.055	0.302	0.026
J (pvalue)		0.184		0.125		0.306		0.597		0.294		0.173		0.197
F_{Eff}		66.654		66.654		66.654		56.579		53.526		66.654		67.799

Notes: OLS represents Ordinary Least Squares and IV-GMM is Instrumental variables Generalized Method of Moments. Instruments are policy shocks and in the coal- (gas-) based generation equation, the gas (coal) price. In the coal generation equation, we report only those results with TTF as the gas price instrument. Effective F (F_{Eff}) is computed based on Montiel Olea and Pflueger (2013). Standard errors are robust for arbitrary forms of serial correlation and heteroskedasticity. For a given gas price variable, the first-stage regression and hence F_{Eff} remains unchanged as the electricity price variables are cycled; that is unless the N is changing.
***p $\leq$ 0.01, **p $\leq$ 0.05, *p $\leq$ 0.1.

of 10, decisively rejecting the null that the instruments are weak (Andrews et al., 2019).

The estimates for technology-specific carbon intensity (i.e., the coefficient on the allowance price) generally differ between OLS and IV. The OLS imputed gas carbon intensities are greater than their IV imputed counterparts. In contrast, the OLS imputed coal carbon intensities are largely in agreement with the IV imputed, but where there is some divergence, the former are negatively biased. A positive correlation between the allowance price and the unobserved mark-up for gas-fired dispatch, implies that OLS overestimates the gas-firing carbon intensity. By contrast, the negative bias in the coal carbon intensity stems from a negative relationship between the allowance price and the mark-up for coal-fired dispatch. Because coal generally sits below gas in the merit order, greater correlation between gas-fired dispatch and its unobserved mark-up implies greater bias in the gas dispatch equation (than in the coal dispatch equation).

To infer pass-through, we compare the empirically imputed technology-specific carbon intensities to the engineering technology-specific carbon intensities of Table 1. Due to the large standard errors on the empirically imputed IV carbon intensities, and with the exception of France, a formal statistical test does not reject the equivalence between these coefficients.[21] These results thus imply that pass-through is complete for most markets and technologies, and in case of the French market more than complete. Moreover, because these imputed intensities are statistically different than zero—mostly at the 1% level of significance—we can also conclude that allowance costs bite.

The significance of the IV estimates in Table 6 lies not just in being able to impute pass-through but in, for the first time, making technology-specific inference. Take Germany's gas-fired generation for instance. The IV results using TTF gas futures price predict an increase in the electricity price of €0.457/MWh for a €1/tCO2 change in the allowance price. The NBP (Zeebrugge) gas futures price predicts slightly higher pass-through €0.582/MWh (€0.617/MWh). For coal, the IV results predict an increase in the electricity price of €0.968/MWh for a €1/tCO2 change in the allowance price. These figures compare well with the ≈ 0.4 (≈ 0.9) that is implied by the engineering

[21] The Belgian NBP gas futures carbon intensity estimate is off for reasons we failed to pinpoint. We do not test this entry.

estimates. The estimates for the other countries mostly tell the same story, again with the exception of France. France is a unique market. Unlike the others, it sources nearly 90% of its energy from nuclear and half the time nuclear is the marginal technology. Small but potentially crucial roles for gas and coal in the French electricity grid may explain this high pass-through for France.

Besides the results in Table 6 being consistent with complete pass-through, we can also conclude that gas-fired generators pass-through the cost of an allowance at one-third to one-half the magnitude for coal-fired generation. This is what the engineering carbon intensities, when fed into the theory, prescribe.

Correlation between fuel prices and the unobserved technology-specific mark-up implies the imputed heat-rate (i.e., the coefficient on both the gas and coal price) are in general biased and inconsistent. The degree of bias for these varies, however. Gas-fired generation, which often sits above coal-fired generation in the merit order, has greater correlation with the mark-up. Its imputed heat-rate is, therefore, less reliable. A negative correlation of fuel prices with the unobserved mark-up means that the imputed heat-rates are downward biased. This is especially true for gas where engineering heat-rates are between 1.7 and 2.2 (Table 1), and yet the imputed heat-rates (Table 6) are, with the exception of France, far below this range. The imputed heat-rates for coal though potentially underestimated, are in several cases closer to the engineering heat-rates. This is thanks to coal normally being dispatched first and is thus less correlated with the unobserved mark-up. Gas-based generators thus appear to (at the very least) pass-through at least half of their fuel costs, while coal-based generators appear to fully pass-through fuel costs.

Is allowance cost pass-through stable across different sub-periods? Results provided in Appendix F indicate that pass-through is generally complete in Phase II and that pass-through is greater in Phase III. Those findings must be interpreted with caution, however, since identification depends crucially on the presence of sufficient exogenous variation, yet splitting the sample reduces such variation. Dropping policy shocks as an instrument, and retaining only the substitutes' fuel price, affects the magnitude and precision of our imputed carbon intensities. While these results—also presented in

Appendix F—still indicate an improvement over OLS in being able to correctly impute pass-through, they also overestimate pass-through relative to the case where policy shocks are available as an additional instrument. Another potential robustness check is to use policy shocks as the only instrument. Such an exercise is uninformative for pass-through as the necessary exclusion restriction does not hold.

6 Conclusions

The MSR was appended to the EU ETS, to help dynamically regulate the availability of allowances and promote allowance price stability. It remains unclear, however, to what extent the EU ETS was operating in line with expectations prior to this reform. We set up a parsimonious structural model of a power generator engaged in fuel-switching and allowance banking for emissions compliance. We decomposed the model and showed that in equilibrium: (i) there exists a tight link between the allowance price, and the gas and coal price as its fundamental drivers; (ii) incentives for allowance banking, reducing and smoothing compliance costs, are decomposable from allowance price variations; and (iii) gas, coal, allowance, and electricity prices have a structurally quantifiable relationship, that is instrumental for inferring the pass-through of allowance and fuel costs.

We took the decomposed model to data. Empirical findings were mostly inline with the predictions of the theory. That is, despite its woes, the EU ETS in fact had success in inducing a typical generator in the power industry to internalise carbon costs. In particular, we found—contrary to previous research—that the allowance price responds to variations in both the natural gas and coal price, and gave precise reasoning of why this finding has eluded other authors. We also documented that variations in the allowance price exhibit a statistically significant risk component, suggesting that banked allowances help reduce long-run compliance costs by smoothing compliance cost shocks. Finally, we found that emission costs are fully passed-through by the representative power generator across seven major European power markets. As a first, we imputed pass-through by dispatch technology.

Some issues that our research was unable to address include the long-term (i.e. levels) relationship between allowance prices and the fundamental drivers, as well as long-term emission reductions achievable via rising allowance prices. We leave this task for future work.

Acknowledgements

Marina Friedrich, Beat Hintermann, Michael Pahle, Oliver Tietjen and audiences at PIK, EAERE, and AERE provided valuable comments. Any arising errors are the sole responsibility of the authors. Samuel J. Okullo acknowledges funding from Bundesministerium für Bildung und Forschung (BMBF) under the research project ENavi (Project number: 03SFK5A & 03SFK4B1).

References

Aatola, P., Ollikainen, M., Toppinen, A., 2013. Price determination in the {EU} {ETS} market: Theory and econometric analysis with market fundamentals. Energy Economics 36, 380 – 395.

Andrews, I., Stock, J.H., Sun, L., 2019. Weak instruments in instrumental variables regression: Theory and practice. Annual Review of Economics 11, 727–753. https://doi.org/10.1146/annurev-economics-080218-025643.

Bayer, P., Aklin, M., 2020. The european union emissions trading system reduced co2 emissions despite low prices. Proceedings of the National Academy of Sciences 117, 8804–8812.

Benz, E., Trück, S., 2009. Modeling the price dynamics of co2 emission allowances. Energy Economics 31, 4–15.

Chang, F.R., 2004. Stochastic Optimization in Continuous Time. Cambridge University Press.

Chevallier, J., Sévi, B., 2014. On the stochastic properties of carbon futures prices. Environmental and Resource Economics 58, 127–153.

Conrad, C., Rittler, D., Rotfuß, W., 2012. Modeling and explaining the dynamics of european union allowance prices at high-frequency. Energy Economics 34, 316–326.

Creti, A., Jouvet, P.A., Mignon, V., 2012. Carbon price drivers: Phase i versus phase {II} equilibrium? Energy Economics 34, 327 – 334.

ENERDATA, 2018. Efficiency of gas- and coal- fired.

Fabra, N., Reguant, M., 2014. Pass-through of emissions costs in electricity markets. American Economic Review 104, 2872–99.

Fan, Y., Jia, J.J., Wang, X., Xu, J.H., 2017. What policy adjustments in the eu ets truly affected the carbon prices? Energy Policy 103, 145–164.

Fell, H., Hintermann, B., Vollebergh, H., 2015. Carbon content of electricity futures in phase ii of the eu ets. Energy Journal 36, 61–83.

French, K.R., Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. Journal of financial Economics 19, 3–29.

Fuss, S., Flachsland, C., Koch, N., Kornek, U., Knopf, B., Edenhofer, O., 2018. A framework for assessing the performance of cap-and-trade systems: Insights from the european union emissions trading system. Review of Environmental Economics and Policy 12, 220–241.

Germeshausen, R., 2020. The european union emissions trading scheme and fuel efficiency of fossil fuel power plants in germany. Journal of the Association of Environmental and Resource Economists 7, 751–777. https://doi.org/10.1086/708894.

Ghysels, E., Santa-Clara, P., Valkanov, R., 2005. There is a risk-return trade-off after all. Journal of Financial Economics 76, 509–548.

Ghysels, E., Santa-Clara, P., Valkanov, R., 2006. Predicting volatility: getting the most out of return data sampled at different frequencies. Journal of Econometrics 131, 59–95.

Guo, B., Gissey, G.C., 2021. Cost pass-through in the british wholesale electricity market. Energy Economics 102, 105497.

Heindl, P., 2017. The impact of administrative transaction costs in the eu emissions trading system. Climate Policy 17, 314–329. http://dx.doi.org/10.1080/14693062.2015.1110108.

Hintermann, B., 2010. Allowance price drivers in the first phase of the {EU} {ETS}. Journal of Environmental Economics and Management 59, 43 – 56.

Hintermann, B., 2016. Pass-through of co2 emission costs to hourly electricity prices in germany. Journal of the Association of Environmental and Resource Economists 3, 857–891. https://doi.org/10.1086/688486.

Hintermann, B., Peterson, S., Rickels, W., 2016. Price and market behavior in phase ii of the eu ets: A review of the literature. Review of Environmental Economics and Policy 10, 108–128.

Jaraitė, J., Convery, F., Di Maria, C., 2010. Transaction costs for firms in the eu ets: lessons from ireland. Climate Policy 10, 190–215.

Kennedy, P.E., 2005. Oh No! I Got the Wrong Sign! What Should I Do? The Journal of Economic Education 36, 77–92.

Kling, C., Rubin, J., 1997. Bankable permits for the control of environmental pollution. Journal of Public Economics 64, 101–115.

Koch, N., 2014. Dynamic linkages among carbon, energy and financial markets: A smooth transition approach. Applied Economics 46, 715–729.

Koch, N., Fuss, S., Grosjean, G., Edenhofer, O., 2014. Causes of the eu ets price drop: Recession, cdm, renewable policies or a bit of everything? — new evidence. Energy Policy 73, 676 – 685.

Koch, N., Grosjean, G., Fuss, S., Edenhofer, O., 2016. Politics matters: Regulatory events as catalysts for price formation under cap-and-trade. Journal of Environmental Economics and Management 78, 121 – 139.

Li, Q., Racine, J.S., 2007. Nonparametric econometrics: theory and practice. Princeton University Press.

Lutz, B.J., Pigorsch, U., Rotfuß, W., 2013. Nonlinearity in cap-and-trade systems: The eua price and its fundamentals. Energy Economics 40, 222–232.

Mansanetbataller, M., Pardo, A., Valor, E., 2007. Co2 prices, energy and weather. The Energy Journal 28, 73–92.

Merton, R.C., 1973. An intertemporal capital asset pricing model. Econometrica: Journal of the Econometric Society , 867–887.

Montiel Olea, J.L., Pflueger, C., 2013. A robust test for weak instruments. Journal of Business & Economic Statistics 31, 358–369.

Paolella, M.S., Taschini, L., 2008. An econometric analysis of emission allowance prices. Journal of Banking & Finance 32, 2022–2032.

Rickels, W., Görlich, D., Peterson, S., 2015. Explaining european emission allowance price dynamics: Evidence from phase ii. German Economic Review 16, 181–202.

Rousseeuw, P.J., Leroy, A.M., 2005. Robust regression and outlier detection. volume 589. John wiley & sons.

RTE, 2018. Efficiency of gas- and coal- fired.

Rubin, J.D., 1996. A model of intertemporal emission trading, banking, and borrowing. Journal of Environmental Economics and Management 31, 269 – 286.

Schennach, S.M., 2000. The economics of pollution permit banking in the context of title iv of the 1990 clean air act amendments. Journal of Environmental Economics and Management 40, 189–210.

Sijm, J., Hers, J., Lise, W., Wetzelaer, B., 2008. The impact of the EU ETS on electricity prices. Final report to DG Environment of the European Commission. Technical Report. Energy research Centre of the Netherlands ECN.

Sijm, J., Neuhoff, K., Chen, Y., 2006. Co2 cost pass-through and windfall profits in the power sector. Climate policy 6, 49–72.

Verardi, V., Croux, C., 2009. Robust regression in stata. The Stata Journal 9, 439–453.

Wooldridge, J.M., 2016. Introductory Econometrics. Cengage learning. sixth edition.

Wooldridge, J.M., 2018. Introductory Econometrics. Cengage learning. seventh edition.

A Notable events in the history of the EU ETS

Table A.1: Policy events in Phase II and III of the EU ETS.

Date	Event
8-Oct-08	Regulation 994/2008 for a standardized and secured registry.
19-Nov-08	Directive 2008/101/EC amending Directive 2003/87/EC so as to include aviation activities.
Dec 12, 2008	Parliament achieves agreement with council on all proposals of the package
April 6, 2009	Council adopts legislative package with 20% target and linear reduction factor of 1.74%
23-Apr-09	Directive 2009/29/EC amending Directive 2003/87/EC to improve and extend the EU ETS.
Mar 29, 2011	Commission announces proposed "Energy Efficiency Directive."
May 30, June 30, 2011	German moratorium and curtailment of nuclear in the power system in response to Fukushima.
Dec 20, 2011	EP ENVI committee passes amendment of energy efficiency directive including withdrawal of 1.4 billion EUA
Feb 28, 2012	EP ITRE committee adopts amendment of EE directive and calls for report on set-aside plan
June 13/14, 2012	Energy Efficiency Directive agreed to as directive 2012/27/EU.
Nov 14, 2012	European Commission tables a draft amendment to postpone the auctioning of 900 million allowances.
Oct 25, 2012	Energy efficiency "Directive 2012/27/EU" adopted.
Jan 24, 2013	Down vote by European Parliament Energy and Industry Committee on back-loading proposal.
Feb 19, 2013	European Parliament's vote to favour an amendment in favour of back-loading.
Apr 16, 2013	Back-loading proposal voted down by Members of the European Parliament (MEPs).
Jul 3, 2013	Second vote on back-loading proposal passes in European parliament.
Jan 22, 2014	European Commission proposes the introduction of the MSR to improve performance of the EU ETS starting 2021.
Feb 27, 2014	Back-loading amendment of the EU ETS auctioning regulation enters into force.
Apr 2014	Back-loading of 900 million allowances starts.
Jul 23, 2014	Amendments to "Directive 2012/27/EU" proposed.
Feb 12, 2015	Senior Members of the European Parliament vote for earlier MSR fix to the EU ETS.
Feb 24, 2015	Members of the European Parliament vote for earlier MSR fix to the EU ETS.
March 6, 2015	Nationally determined reduction commitment of the Union and its Member States submitted to the UNFCCC.
May 13, 2015	Members states permanent representative endorse information MSR agreement between Council and Parliament.
July 8, 2015	European parliament approves EU ETS market stability reserve.
Jul 15, 2015	EC presents its original proposal on EU ETS reform.
Oct 6, 2015	Decision EU 2015/1814 of the European Council and Parliament establishes the MSR.
Dec 12 2015	21st Conference of the Parties (COP 21) concludes with adoption of the Paris Agreement.
May 12, 2016	German minister Sigmar Gabriel rebuffs French proposal for an EU carbon price corridor.
Oct 5, 2016	EU parliament approves ratification of Paris agreement.
Nov 4, 2016	The Paris Agreement, adopted on 12 December 2015 enters into force.
Feb 12, 2017	European Parliament and Council reach their respective positions on the EU ETS reform.
Feb 16, 2017	European parliament agrees to reduce the ceiling by 2.2.% after 2020 and double intake capacity of MSR.
Feb 28, 2017	European Union Ministers reach agreement on reform of EU ETS.
May 12, 2017	Commission publishes first surplus indicator of EU ETS MSR.
Nov 10, 2017	European law makers provisionally agree on the reforms of the EU ETS for the period after 2020.
Nov, 22 2017	Press release: reform of the EU emissions trading system, Council endorses deal with European Parliament.
Feb 27, 2018	Council of the European Union approves reform of the EU ETS for the period after 2020.
March 14, 2018	Directive (EU) 2018/410 of the European Parliament and of the Council amending Directive 2003/87/EC to enhance cost-effective emission reductions and low-carbon investments, and Decision (EU) 2015/1814.
Dec 11, 2018	"Directive 2018/2002/EU" amending "Directive 2012/27/EU" on energy efficiency adopted.

Notes: Hand collected from the Financial Times, Google Search, and the European Commission Climate Action new archive.

B Summary statistics

Table A.2: Summary statistics (2008-2018) based on weekly averages of daily data. See A.3 for summary statistics based on daily data.

Variable	Level					Growth rate				
	Obs.	Mean	Std. Dev	Min	Max	Obs.	Mean	Std. Dev	Min	Max
Allowance price €/tCO2										
month-ahead future	539	10.18	5.86	2.94	30.33	539	-0.09	5.54	-43.01	15.77
year-ahead future	537	10.06	5.58	2.99	28.07	537	-0.10	5.56	-42.84	16.07
spot	527	8.84	4.68	0.01	24.78	526	-0.24	6.89	-62.86	40.55
Coal (€/MWh)										
Coal API2	539	8.32	2.18	4.79	16.84	539	-0.04	3.04	-15.41	18.01
ICE Rotterdam year-ahead	538	8.46	2.03	4.08	16.31	538	-0.02	3.11	-17.24	14.73
Gas (€/MWh)										
ICE TTF month-ahead	539	20.30	5.38	8.45	35.21	539	-0.03	4.12	-29.31	14.29
ICE TTF year-ahead	539	22.29	5.38	12.23	41.82	539	-0.04	3.27	-24.53	25.15
ICE Zebrugge month-ahead	539	20.29	5.48	8.27	35.47	539	-0.03	4.49	-33.09	22.24
ICE Zebrugge year-ahead	539	22.60	5.59	12.39	43.22	539	-0.03	3.20	-13.32	29.65
ICE NBP month-ahead	539	20.36	5.48	8.08	36.62	539	-0.03	4.75	-32.01	24.55
ICE NBP year-ahead	539	22.76	5.55	11.97	43.20	539	-0.03	3.58	-20.84	29.70
Electricity (baseload month-ahead, €/MWh)										
EEX Germany	539	41.96	12.85	21.33	94.98	539	-0.13	4.32	-14.62	17.13
Dutch	539	46.76	12.54	24.29	102.21	539	-0.09	4.05	-16.27	15.90
ICE UK	539	47.49	12.17	31.36	117.10	539	-0.02	3.80	-19.93	24.64
Belgium	506	46.90	13.05	22.80	116.10	505	-0.12	6.31	-24.01	28.93
Italy	496	60.08	12.28	31.70	101.28	495	-0.12	3.66	-16.10	19.90
France	539	47.33	15.70	21.53	117.17	539	-0.16	7.19	-41.53	31.30
Nordpool	538	41.89	12.80	21.33	95.38	537	-0.13	4.24	-14.89	16.85
Controls										
EURO STOXX 200	539	316.49	55.09	175.51	426.34	539	0.00	2.16	-12.68	6.31
Norway Water Reservoir % Level Area	539	60.98	19.69	16.70	91.40	539	-0.21	6.02	-13.27	34.16

Notes: All data except allowance prices sourced from Thompson Rueters datastream. Allowance prices obtained from the Inter-Continental Exchange through "www.quandl.com" and spot allowance prices from the European Energy Exchange.

Table A.3: Summary statistics (2008-2018). Note that these are presented for daily frequency data. The analysis is conducted using the weekly frequency.

Variable	Level					Growth rate				
	Obs.	Mean	Std. Dev	Min	Max	Obs.	Mean	Std. Dev	Min	Max
Allowance price €/tCO2										
month-ahead future	2,702	10.18	5.86	2.70	31.71	2,702	-0.02	3.11	-43.47	23.92
year-ahead future	2,690	10.06	5.58	2.75	29.33	2,690	-0.02	3.11	-43.21	24.52
spot	2,568	8.86	4.66	0.01	25.00	2,524	-0.03	6.29	-69.31	69.31
Coal (€/MWh)										
Coal API2	2,702	8.32	2.19	4.72	17.46	2,702	-0.01	1.62	-21.98	16.38
ICE Rotterdam year-ahead	2,699	8.46	2.03	3.64	16.95	2,699	0.00	1.60	-20.31	14.39
Gas (€/MWh)										
ICE TTF month-ahead	2,702	20.30	5.40	7.84	36.15	2,702	-0.01	2.30	-11.65	26.72
ICE TTF year-ahead	2,702	22.29	5.40	11.69	42.12	2,702	-0.01	2.96	-79.27	82.91
ICE Zebrugge month-ahead	2,702	20.29	5.50	7.43	39.98	2,702	-0.01	2.98	-27.39	27.74
ICE Zebrugge year-ahead	2,702	22.59	5.59	12.03	43.62	2,702	-0.01	1.93	-42.52	45.51
ICE NBP month-ahead	2,702	20.37	5.50	7.08	39.14	2,702	-0.01	2.90	-16.70	31.63
ICE NBP year-ahead	2,702	22.75	5.56	11.77	43.68	2,702	-0.01	3.39	-102.54	101.98
Electricity (baseload month-ahead, €/MWh)										
EEX Germany	2,702	41.98	12.88	21.11	98.41	2,702	-0.03	2.38	-18.75	18.88
Dutch	2,702	46.77	12.56	24.05	104.63	2,702	-0.02	2.12	-15.63	17.82
ICE UK	2,702	47.50	12.19	31.31	122.25	2,702	0.00	1.85	-18.05	20.47
Belgium	2,533	46.87	13.12	22.58	122.25	2,532	-0.02	5.04	-106.02	104.81
Italy	2,483	60.04	12.22	31.50	102.25	2,482	-0.02	1.94	-21.00	17.48
France	2,702	47.37	15.83	21.23	138.00	2,702	-0.03	3.78	-42.96	42.39
Nordpool	2,694	41.88	12.80	21.13	98.13	2,693	-0.02	2.37	-30.85	18.66
Controls										
EURO STOXX 200	2,702	316.39	55.15	170.80	428.09	2,702	0.00	1.27	-8.18	9.82
Norway Water Reservoir % Level Area	2,702	61.03	19.71	6.00	91.40	2,702	-0.04	7.13	-237.65	244.09

Notes: All data except allowance prices sourced from Thompson Rueters datastream. Allowance prices obtained from the Inter-Continental Exchange through "www.quandl.com" and spot allowance prices from the European Energy Exchange.

C Additional graphs

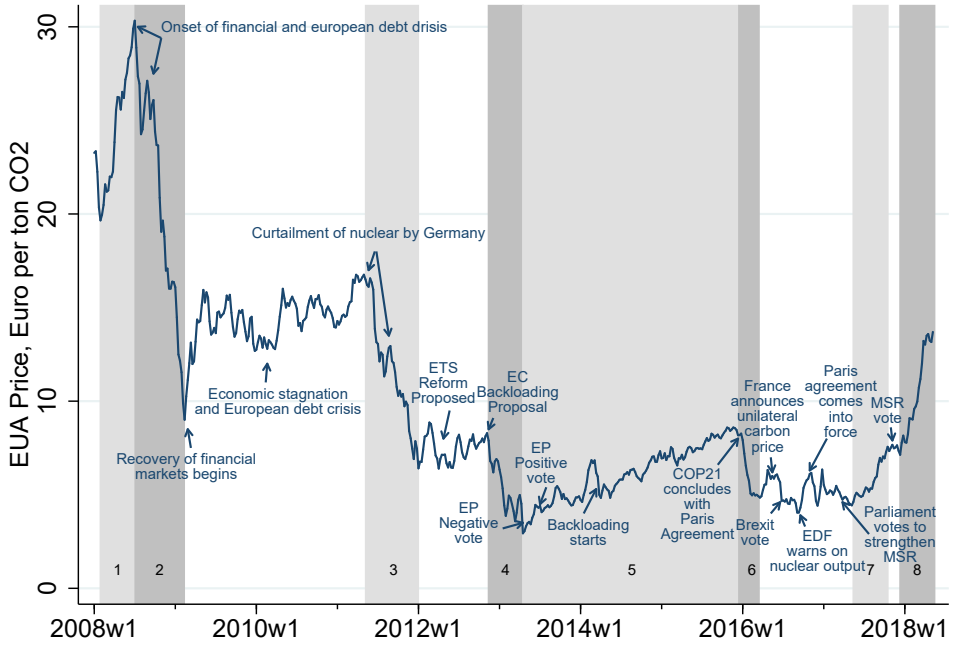


Figure A.1: Evolution of the EU ETS allowance price, some historical events, and regimes for notable structural shifts in the trend of the allowance that were spurred by a policy disclosure directly relating to the EU ETS and macroeconomic shocks.

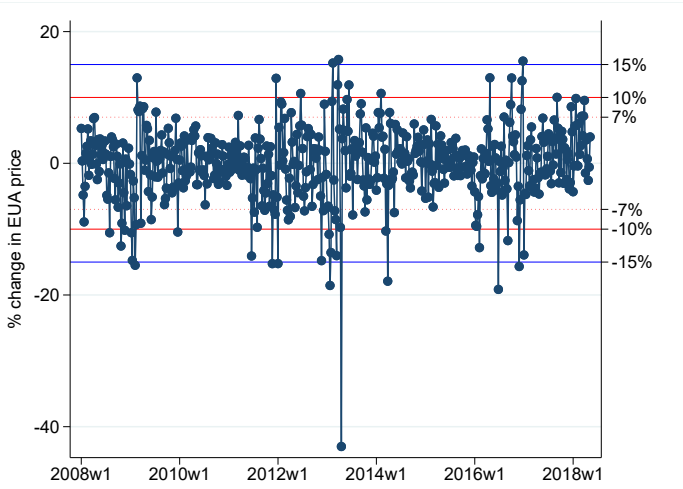


Figure A.2: Percent change in the EUA price.

Table A.4: Elasticity of EUA month-ahead futures to natural gas and coal month-ahead futures.

	Short-run propensities								Long-run propensities	
Covariate	(1) Static	(2) Dynamic	(3) Robust	(4) Non par.	(5) Static	(6) Dynamic	(7) Robust	(8) Non par.	(9)	(10)
Natural gas	0.184***	0.165***	0.216***	0.144***					0.184***	
	(0.056)	(0.052)	(0.049)	(0.043)					(0.058)	
Coal	-0.209***	-0.190***	-0.199***	-0.158***					-0.211***	
	(0.078)	(0.068)	(0.067)	(0.061)					(0.077)	
Relative coal price					-0.191***	-0.172***	-0.212***	-0.150***		-0.191***
					(0.053)	(0.048)	(0.045)	(0.045)		(0.054)
Auto-regressive lag	No	Yes	No	No	No	Yes	No	No	No	No
N	539	539	539	539	539	539	539	539	539	539
adj. R-sq	0.487	0.496	0.478	0.498	0.488	0.497	0.476	0.540	0.570	0.570

Notes: Coefficients in (1)-(8) are impact propensities whereas those in specifications (9) and (10) are long-run propensities (Wooldridge, 2018, p. 607) calculated using specifications (2) and (6). Standard errors for (1) and (5) are New-West heteroskedasticity and serial autocorrelation consistent standard errors. Those in (2) and (6), where the serial correlation is modelled explicitly, are White's heteroskedasticity robust standard errors. One lag of EUA returns proves sufficient in modelling the serial autocorrelation present in EUA price returns. After modelling the serial correlation, we find that the null that errors are homoskedastic cannot be rejected; nonetheless, for conservativeness we report heteroskedasticity robust standard errors. Specification (3) and (7) eliminate outliers based on Cook's distance, and Huber and biweight iterations. Specifications (4) and (8) are based on a non-parametric kernel regression which is robust to functional form misspecification. All regressions are based on the 3-Jan-2008 to 9-May-2019 sample. Week, month, and year dummies are used as controls in all regressions. Dummies are also introduced to control for periods of systematic trending in the allowance price. All specifications except (3) and (7) include dummies to control for week over week returns in excess of 10%. Such drastic movements in the allowance occur only occasionally, and yet can have substantial influence on estimated standard errors. These significant movements can one-for-one be linked to regulatory announcements.
***Significant at 1 percent.

D Additional results for dependence of allowance prices on fundamentals

E Additional results for the incentive of banking the marginal allowance

Table A.5: The responsiveness of EUA returns to the scarcity and volatility premium.

	Phase II-III			Phase II			Phase III		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
MiDaS									
Discount rate		-5.421	-0.878		-2.380	-0.367		-0.130	-0.870
		(3.549)	(3.488)		(1.722)	(1.632)		(4.184)	(4.064)
Volatility	11.476***	11.667***	11.54***	9.405	10.263	10.671	11.303**	11.161**	11.02**
	(4.226)	(4.198)	(4.216)	(7.052)	(7.025)	(7.098)	(5.464)	(5.433)	(5.432)
N	539	539	539	260	260	260	279	279	279
adj. R-sq	0.537	0.505	0.492	0.492	0.460	0.437	0.551	0.545	0.539
BIC	3403.469	3402.087	3405.610	1706.972	1705.296	1708.058	1888.248	1887.439	1882.133

Notes: Volatility is predicted using daily absolute returns (Ghysels et al., 2006). Column (1) uses the allowance price returns net of the observed risk-free rate as the predictor variable. Column (2) and (3) both use unadjusted returns as the predicted variable. Because column (1) uses excess returns, the constant in this equation is not given the interpretation of a risk-free discount rate and is therefore not reported. Relative to Column (2), Column (3) excludes gas and coal price returns as predictors of returns. Volatility is computed using the step-wise MiDaS, where the duration of a step is the number of trading days in the week—five in the absence of a holiday. Two steps are used for predicting conditional volatility. In all regressions, we control for serial correlation using the lagged returns (cf., Ghysels et al., 2005; French et al., 1987). Standard errors are robust for heteroskedasticity. Because the estimating equation (10) is an equilibrium relationship, we make use of 'year' and 'week of year' as additional controls. Moreover, dummies for uncharacteristic changes in EUA prices, and dummies for periods of systematic trending are also used to control for events or outcomes that were unanticipated by the generator at the time the purchasing decisions was taken.
***Significant at 1 percent.

F Additional results for pass-through of allowance costs

Table A.6: Responsiveness of baseload month-ahead electricity futures prices to month-ahead futures gas price and the API2 coal settlement price. Results based on IV-GMM with regulatory shocks used as instruments for the allowance price.

	Electricity price (€/MWh)						
	Germany	Netherlands	U. Kingdom	Belgium	Italy	France	Nordpool
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Title Transfer Facility							
Allowance (€/tCO2)	0.529***	0.461***	0.671***	0.560*	0.547***	0.836***	0.449***
	(0.0903)	(0.135)	(0.125)	(0.255)	(0.155)	(0.168)	(0.107)
Coal (€/MWh)	0.684*	0.926***	1.121***	1.379**	1.078**	1.434**	0.776**
	(0.346)	(0.260)	(0.286)	(0.449)	(0.353)	(0.522)	(0.263)
Gas (€/MWh)	1.073***	1.440***	1.065***	1.402***	0.948***	1.946***	1.069***
	(0.0717)	(0.0759)	(0.0861)	(0.122)	(0.120)	(0.119)	(0.0784)
N	539	539	539	505	495	539	537
J (p value)	0.906	0.761	0.840	0.747	0.748	0.701	0.940
Zeebrugge							
Allowance (€/tCO2)	0.686***	0.644***	0.768***	0.589*	0.620***	1.052***	0.644***
	(0.0907)	(0.134)	(0.139)	(0.284)	(0.137)	(0.165)	(0.101)
Coal (€/MWh)	0.980**	1.182***	1.070**	1.498**	1.255***	1.769***	0.950***
	(0.342)	(0.261)	(0.329)	(0.485)	(0.353)	(0.512)	(0.281)
Gas (€/MWh)	0.963***	1.282***	0.990***	1.126***	0.779***	1.554***	0.937***
	(0.0709)	(0.0494)	(0.0756)	(0.109)	(0.104)	(0.130)	(0.0725)
N	539	539	539	505	495	539	537
J (p value)	0.872	0.813	0.858	0.783	0.765	0.659	0.881
National Balancing Point							
Allowance (€/tCO2)	0.620***	0.506***	0.718***	0.440	0.725***	0.960***	0.595***
	(0.0962)	(0.136)	(0.136)	(0.285)	(0.127)	(0.165)	(0.109)
Coal (€/MWh)	1.300***	1.318***	1.300***	1.677***	1.412***	1.705**	1.181***
	(0.358)	(0.244)	(0.357)	(0.478)	(0.342)	(0.521)	(0.295)
Gas (€/MWh)	0.808***	1.205***	0.894***	1.104***	0.695***	1.490***	0.809***
	(0.0541)	(0.0562)	(0.0806)	(0.130)	(0.0949)	(0.137)	(0.0592)
N	539	539	539	505	495	539	537
J (p value)	0.910	0.821	0.846	0.813	0.771	0.650	0.896

Notes: The electricity price is projected on both the gas and coal price within the same estimating equation. Standard errors are robust for arbitrary forms of serial correlation and heteroskedasticity. Policy shocks instrument for the EUA price.
***p $\leq$ 0.01, **p $\leq$ 0.05, *p $\leq$ 0.1.

Table A.7: Responsiveness of base-load month-ahead electricity futures prices to month-ahead futures gas price and the API2 coal settlement price (Phase 2).

	Electricity price (€/MWh)													
	German Baseload		Dutch Baseload		UK Baseload		Belgian Baseload		IT Baseload		French Baseload		Nordpool Baseload	
	(1) OLS	(2) IV-GMM	(3) OLS	(4) IV-GMM	(5) OLS	(6) IV-GMM	(7) OLS	(8) IV-GMM	(9) OLS	(10) IV-GMM	(11) OLS	(12) IV-GMM	(13) OLS	(14) IV-GMM
PANEL A: natural gas technology (TTF)														
Allowance (€/tCO2)	0.572***	0.574***	0.653***	0.582***	0.888***	0.810***	0.395	0.433#	0.594**	0.589***	0.577*	0.339	0.517**	0.318
	(0.209)	(0.196)	(0.203)	(0.215)	(0.212)	(0.191)	(0.285)	(0.285)	(0.271)	(0.205)	(0.340)	(0.317)	(0.210)	(0.233)
Natural gas (€/MWh)	1.188***	1.151***	1.331***	1.394***	1.455***	1.374***	1.800***	1.526***	0.684***	0.669***	1.740***	1.672***	1.199***	1.193***
	(0.172)	(0.136)	(0.160)	(0.122)	(0.223)	(0.124)	(0.246)	(0.137)	(0.127)	(0.099)	(0.231)	(0.174)	(0.170)	(0.118)
N	260.000	260.000	260.000	260.000	260.000	260.000	226.000	226.000	216.000	216.000	260.000	260.000	258.000	258.000
Adj. R-squared	0.451	0.040	0.488	0.111	0.538	0.225	0.556	0.123	0.390	-0.183	0.527	0.044	0.449	0.044
PANEL B: natural gas technology (NBP)														
Allowance(€/tCO2)	0.675***	0.614***	0.742***	0.544**	0.995***	0.824***	0.396	0.075	0.617**	0.450**	0.728*	0.176	0.616**	0.346
	(0.241)	(0.226)	(0.231)	(0.253)	(0.225)	(0.221)	(0.319)	(0.328)	(0.286)	(0.212)	(0.369)	(0.329)	(0.239)	(0.250)
Natural gas (€/MWh)	0.836***	0.780***	1.080***	1.109***	1.132***	1.145***	1.335***	1.280***	0.509***	0.465***	1.228***	1.297***	0.874***	0.845***
	(0.166)	(0.124)	(0.139)	(0.097)	(0.214)	(0.112)	(0.216)	(0.142)	(0.122)	(0.093)	(0.226)	(0.174)	(0.160)	(0.120)
N	260.000	260.000	260.000	260.000	260.000	260.000	226.000	226.000	216.000	216.000	260.000	260.000	258.000	258.000
Adj R-squared	0.404	-0.043	0.478	0.090	0.508	0.173	0.521	0.059	0.368	-0.231	0.483	-0.059	0.407	-0.032
PANEL C: natural gas technology (ZEE)														
Allowance (€/tCO2)	0.666***	0.595**	0.750***	0.889***	0.996***	1.043***	0.408	0.319	0.625**	0.530**	0.733**	0.487#	0.620***	0.577**
	(0.226)	(0.233)	(0.222)	(0.244)	(0.222)	(0.194)	(0.318)	(0.287)	(0.278)	(0.211)	(0.364)	(0.320)	(0.230)	(0.244)
Natural gas (€/MWh)	1.095***	1.074***	1.285***	1.258***	1.399***	1.299***	1.458***	1.325***	0.559***	0.544***	1.487***	1.475***	1.085***	1.024***
	(0.143)	(0.110)	(0.132)	(0.105)	(0.193)	(0.106)	(0.204)	(0.135)	(0.121)	(0.099)	(0.210)	(0.161)	(0.156)	(0.119)
N	260.000	260.000	260.000	260.000	260.000	260.000	226.000	226.000	216.000	216.000	260.000	260.000	258.000	258.000
Adj R-squared	0.450	0.038	0.506	0.141	0.558	0.258	0.516	0.053	0.379	-0.206	0.505	-0.001	0.442	0.035
PANEL D: coal technology (CIF ARA API2)														
Allowance (€/tCO2)	0.786***	0.943***	0.897***	1.125***	1.145***	1.063***	0.614*	1.027***	0.744**	0.871***	0.909**	1.169***	0.722***	0.884***
	(0.239)	(0.253)	(0.232)	(0.225)	(0.238)	(0.193)	(0.353)	(0.268)	(0.286)	(0.214)	(0.388)	(0.267)	(0.243)	(0.233)
Coal (€/MWh)	1.859**	1.646**	1.922**	2.012***	2.565***	1.661***	3.222***	3.250***	2.143***	2.123***	2.002*	2.918***	1.896**	1.772***
	(0.930)	(0.648)	(0.887)	(0.660)	(0.830)	(0.538)	(0.674)	(0.531)	(0.683)	(0.553)	(1.190)	(0.812)	(0.826)	(0.577)
N	260.000	260.000	260.000	260.000	260.000	260.000	226.000	226.000	216.000	216.000	260.000	260.000	258.000	258.000
Adj R-squared	0.343	-0.152	0.348	-0.136	0.402	-0.024	0.372	-0.238	0.375	-0.214	0.399	-0.227	0.334	-0.153

Notes: OLS represents Ordinary Least Squares and IV-GMM is Instrumental variables Generalized Method of Moments. Instruments are policy shocks and in the coal- (gas-) based generation equation, the gas (coal) price. Effective F computed based on Montiel Olea and Pflueger (2013). Standard errors are robust for arbitrary forms of serial correlation and heteroskedasticity.
***p $\leq$ 0.01, **p $\leq$ 0.05, *p $\leq$ 0.1.

Table A.8: Responsiveness of baseload month-ahead electricity futures prices to month-ahead futures gas price and the API2 coal settlement price (Phase 3).

	Electricity price (€/MWh)													
	German Baseload		Dutch Baseload		UK Baseload		Belgian Baseload		IT Baseload		French Baseload		Nordpool Baseload	
	(1) OLS	(2) IV-GMM	(3) OLS	(4) IV-GMM	(5) OLS	(6) IV-GMM	(7) OLS	(8) IV-GMM	(9) OLS	(10) IV-GMM	(11) OLS	(12) IV-GMM	(13) OLS	(14) IV-GMM
PANEL A: natural gas technology (TTF)														
Allowance (€/tCO2)	0.812***	0.679***	0.271	0.356*	0.230	0.154	-0.448	0.259	0.387#	-0.208	1.121**	1.227**	0.789***	0.700***
	(0.274)	(0.202)	(0.263)	(0.188)	(0.361)	(0.110)	(0.638)	(0.492)	(0.238)	(0.209)	(0.525)	(0.488)	(0.277)	(0.190)
Natural gas (€/MWh)	1.080***	1.125***	1.705***	1.689***	1.105***	1.124***	1.247***	1.234***	1.835***	1.881***	2.485***	2.116***	1.041***	1.109***
	(0.175)	(0.107)	(0.140)	(0.061)	(0.200)	(0.067)	(0.390)	(0.204)	(0.165)	(0.102)	(0.482)	(0.203)	(0.172)	(0.100)
N	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000
Adj. R-squared	0.410	0.120	0.620	0.449	0.207	-0.073	0.230	-0.224	0.606	0.268	0.373	-0.027	0.404	0.100
PANEL B: natural gas technology (NBP)														
Allowance(€/tCO2)	0.907***	0.755***	0.434#	0.348**	0.335	0.295**	-0.345	0.446	0.587**	0.058	1.313**	1.329***	0.881***	0.792***
	(0.273)	(0.207)	(0.264)	(0.166)	(0.374)	(0.118)	(0.621)	(0.500)	(0.251)	(0.192)	(0.525)	(0.473)	(0.274)	(0.196)
Natural gas (€/MWh)	0.929***	0.953***	1.428***	1.399***	0.927***	0.881***	1.089***	0.922***	1.473***	1.448***	2.209***	1.814***	0.894***	0.930***
	(0.136)	(0.084)	(0.118)	(0.063)	(0.181)	(0.054)	(0.341)	(0.176)	(0.129)	(0.079)	(0.420)	(0.171)	(0.136)	(0.080)
N	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000
Adj R-squared	0.412	0.124	0.601	0.422	0.201	-0.080	0.233	-0.223	0.572	0.209	0.384	-0.012	0.405	0.103
PANEL C: natural gas technology (ZEE)														
Allowance (€/tCO2)	0.928***	0.764***	0.442*	0.444***	0.317	0.313***	-0.358	0.276	0.589**	0.002	1.319***	1.362***	0.897***	0.718***
	(0.277)	(0.211)	(0.262)	(0.166)	(0.362)	(0.114)	(0.616)	(0.483)	(0.246)	(0.180)	(0.501)	(0.459)	(0.277)	(0.219)
Natural gas (€/MWh)	0.936***	0.937***	1.512***	1.488***	1.051***	0.954***	1.208***	1.119***	1.575***	1.607***	2.356***	1.988***	0.912***	0.929***
	(0.143)	(0.100)	(0.130)	(0.059)	(0.188)	(0.068)	(0.348)	(0.142)	(0.147)	(0.078)	(0.465)	(0.214)	(0.146)	(0.090)
N	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000
Adj R-squared	0.396	0.100	0.609	0.435	0.225	-0.049	0.238	-0.211	0.582	0.225	0.389	-0.002	0.395	0.086
PANEL D: coal technology (CIF ARA API2)														
Allowance (€/tCO2)	1.213***	1.580***	0.920***	1.922***	0.645*	0.634***	0.026	1.409***	1.090***	1.490***	2.051***	3.720***	1.176***	1.769***
	(0.265)	(0.198)	(0.295)	(0.219)	(0.361)	(0.226)	(0.590)	(0.473)	(0.309)	(0.269)	(0.576)	(0.443)	(0.277)	(0.176)
Coal (€/MWh)	3.240***	2.942***	3.538***	3.614***	2.816**	1.663***	2.616*	2.978***	3.300***	3.713***	6.743***	6.192***	3.133***	3.086***
	(0.645)	(0.413)	(0.545)	(0.368)	(1.188)	(0.507)	(1.375)	(0.624)	(0.575)	(0.451)	(1.499)	(0.989)	(0.636)	(0.364)
N	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000	279.000
Adj R-squared	0.379	0.064	0.339	-0.022	0.149	-0.170	0.203	-0.287	0.407	-0.089	0.331	-0.120	0.376	0.033

Notes: OLS represents Ordinary Least Squares and IV-GMM is Instrumental variables Generalized Method of Moments. Instruments are policy shocks and in the coal- (gas-) based generation equation, the gas (coal) price. Effective F computed based on Montiel Olea and Pflueger (2013). Standard errors are robust for arbitrary forms of serial correlation and heteroskedasticity.
***p $\leq$ 0.01, **p $\leq$ 0.05, *p $\leq$ 0.1.

Table A.9: Responsiveness of baseload month-ahead electricity futures prices to month-ahead futures gas price and the API2 coal settlement price (with only the substitute's fuel costs as the instrument).

	Electricity price (€/MWh)													
	German Baseload		Dutch Baseload		UK Baseload		Belgian Baseload		IT Baseload		French Baseload		Nordpool Baseload	
	(1) OLS	(2) IV-GMM	(3) OLS	(4) IV-GMM	(5) OLS	(6) IV-GMM	(7) OLS	(8) IV-GMM	(9) OLS	(10) IV-GMM	(11) OLS	(12) IV-GMM	(13) OLS	(14) IV-GMM
PANEL A: natural gas technology (TTF)														
Allowance (€/tCO2)	0.660***	0.531*	0.608***	0.512**	0.830***	0.588*	0.201	0.595	0.582***	0.574**	0.654***	0.689	0.621***	0.524*
	(0.168)	(0.310)	(0.165)	(0.260)	(0.208)	(0.305)	(0.251)	(0.364)	(0.182)	(0.253)	(0.247)	(0.427)	(0.171)	(0.305)
Natural gas (€/MWh)	1.174***	1.190***	1.514***	1.526***	1.309***	1.338***	1.607***	1.584***	1.111***	1.112***	2.026***	2.022***	1.166***	1.178***
	(0.128)	(0.123)	(0.116)	(0.107)	(0.188)	(0.165)	(0.197)	(0.183)	(0.149)	(0.137)	(0.258)	(0.245)	(0.127)	(0.121)
N	539	539	539	539	539	539	505	505	495	495	539	539	537	537
Adj. R-squared	0.420	0.184	0.492	0.314	0.386	0.201	0.363	0.031	0.456	0.118	0.428	0.117	0.417	0.186
PANEL B: natural gas technology (NBP)														
Allowance(€/tCO2)	0.760***	0.488	0.715***	0.452*	0.931***	0.561**	0.231	0.342	0.628***	0.539**	0.817***	0.571	0.718***	0.478
	(0.190)	(0.347)	(0.188)	(0.273)	(0.220)	(0.275)	(0.256)	(0.443)	(0.196)	(0.264)	(0.267)	(0.477)	(0.192)	(0.342)
Natural gas (€/MWh)	0.909***	0.931***	1.263***	1.283***	1.055***	1.084***	1.268***	1.263***	0.883***	0.890***	1.602***	1.621***	0.919***	0.937***
	(0.113)	(0.112)	(0.103)	(0.100)	(0.156)	(0.149)	(0.175)	(0.165)	(0.132)	(0.120)	(0.243)	(0.234)	(0.107)	(0.106)
N	539	539	539	539	539	539	505	505	495	495	539	539	537	537
Adj R-squared	0.392	0.141	0.486	0.302	0.369	0.174	0.354	0.021	0.430	0.077	0.410	0.088	0.394	0.150
PANEL C: natural gas technology (ZEE)														
Allowance (€/tCO2)	0.770***	0.648**	0.742***	0.663**	0.941***	0.728**	0.252	0.597	0.653***	0.607**	0.848***	0.850**	0.734***	0.634**
	(0.185)	(0.318)	(0.183)	(0.260)	(0.212)	(0.298)	(0.257)	(0.380)	(0.191)	(0.273)	(0.259)	(0.427)	(0.189)	(0.314)
Natural gas (€/MWh)	1.057***	1.066***	1.398***	1.403***	1.234***	1.250***	1.360***	1.346***	0.911***	0.914***	1.792***	1.791***	1.047***	1.054***
	(0.109)	(0.107)	(0.102)	(0.097)	(0.163)	(0.150)	(0.180)	(0.166)	(0.137)	(0.126)	(0.258)	(0.241)	(0.113)	(0.110)
N	539	539	539	539	539	539	505	505	495	495	539	539	537	537
Adj R-squared	0.417	0.180	0.502	0.327	0.402	0.222	0.354	0.017	0.435	0.085	0.420	0.105	0.414	0.181
PANEL D: coal technology (CIF ARA API2)														
Allowance (€/tCO2)	0.913***	1.126***	0.945***	1.282***	1.104***	1.253***	0.507*	1.438***	0.882***	1.212***	1.108***	1.710***	0.871***	1.108***
	(0.192)	(0.339)	(0.192)	(0.282)	(0.197)	(0.279)	(0.277)	(0.408)	(0.199)	(0.297)	(0.286)	(0.448)	(0.201)	(0.332)
Coal (€/MWh)	1.997***	1.973***	2.214***	2.176***	2.491***	2.474***	2.888***	3.009***	2.412***	2.435***	2.881***	2.813***	1.982***	1.955***
	(0.704)	(0.676)	(0.670)	(0.646)	(0.711)	(0.651)	(0.641)	(0.631)	(0.420)	(0.387)	(1.034)	(0.996)	(0.636)	(0.611)
N	539	539	539	539	539	539	505	505	495	495	539	539	537	537
Adj R-squared	0.308	0.024	0.287	0.030	0.281	0.068	0.278	-0.121	0.379	-0.015	0.319	-0.060	0.302	0.021

Notes: OLS represents Ordinary Least Squares and IV-GMM is Instrumental variables Generalized Method of Moments. Instruments are policy shocks and in the coal- (gas-) based generation equation, the gas (coal) price. Effective F computed based on Montiel Olea and Pflueger (2013). Standard errors are robust for arbitrary forms of serial correlation and heteroskedasticity.
***p $\leq$ 0.01, **p $\leq$ 0.05, *p $\leq$ 0.1.

Environmental, Social, and Governance Factors Within Investing: Impact on the Financial Performance of the Energy Sector

ABSTRACT

The purpose of this study is to analyze the effects of ESG activity and ratings on the financial performance of firms in the energy sector comparing renewable and nonrenewable energy companies. Using Timeseries ESG data of the first quarter of 2018 from Morgan Stanley Capital International (MSCI), we create four different portfolios using a sample of 78 energy companies. The portfolios are split by renewable and nonrenewable companies, then further split into four categories based on leading ESG scores and lagging ESG scores. After calculating the holding period return and running the Capital Asset Pricing Model with each portfolio, we find that the laggard ESG nonrenewable energy portfolio generates the best return. Further findings suggest that the high market capitalization, the Covid-19 pandemic, and the short period of observation all play important roles in financial performance.

INTRODUCTION

Every day, market analysts and investors look for any type of pattern which they can use to their advantage to make a profit through the stock market. Valuations are completed to predict future prices, ratios are analyzed, and investment strategies are adjusted. In the past few decades, a new factor has been thrown in the mix to once again adjust how investments are made. With a rising concern regarding environmental stability and overall corporate responsibility, the concept of environmental, social, and governance (ESG) factors has made it to Wall Street and beyond. Essentially, this is the idea of leading sustainable and responsible investment strategies (Dalal, 2019). In turn, firms must model these ethical standards to be viewed in a positive light by investors. This includes leading environmentally friendly business practices, adhering to social standards by treating their employees and the communities around them well, and following governmental regulations as they are set out (Auer & Schumacher, 2015). The incorporation of ESG into investment strategies varies by the firm/investor. Some put it on the forefront, while others find it less important in the big picture. To better understand the extent to which ESG factors effect portfolio performance, this study focused solely on Morgan Stanley Capital International ESG scores to determine portfolio construction with a focus in the energy sector. While it cannot be a sole predictor of investment returns, my research shows both its value and variability from 2018-2021.

Defining Environmental, Social, and Governance

The idea of ESG was first mentioned in 2006 in the United Nations Principles for Responsible Investment report, discussing how it must be incorporated in financial evaluations of firms to influence sustainable investments (Atkins 2020). Since then the research surrounding it within the finance, accounting, and management world has skyrocketed. Essentially, to be a socially responsible investor you must be aware of the firms ESG activity and have your own measurement of quality. To address this issue, there has been an emergence of rating companies which all use their own empirical analysis strategies to determine a final ESG score or rating. Popular rating firms include Morgan Stanley Capital International (MSCI), Sustainalytics, S&P Global, Vigeo-Eiris, or Thomson Reuters Refinitiv ESG (Berg 2021). While the approach to ratings will vary by firm, the areas looked at remain the same. Environmental scores prioritize environmental awareness, proactivity in recycling, waste production, environmental cleanup,

renewable energies, and biotechnology. Social scores reflect issues surrounding labor relations and conditions such as empowerment, employment of minorities, profit sharing, and many more related ideas. Finally, governance scores involve executive compensation, voting and shareholder rights, board independence and elections, along with auditor independence (Auer & Schumacher 2015).

While rating firms are a good influence on sustainable investments, it still raises cconcern for investors and policymakers. Studies have argued that ESG scores have a strong correlation with firm size. Larger firms can provide more data and resources for ESG rating agencies, which then leads to them having higher sustainability scores compared to small firms (Drempetic, 2019). On top of this, since the approaches to ratings are so different between agencies, there is a lot of inconsistency on the evaluation of a firms ESG performance (Atkins 2020). Survey evidence from Amel- Zadeh and Serafeim (2018) shows "82% of investment professionals use ESG information in the investment process, but 26.4% also indicate a lack of ESG rating reliability" (Atkins 2020). This means that the same firm might have two completely different evaluations depending on what rating agency one is looking at. With a lack of consistency in scores, it forces investors to choose what agency they base analyzations off. This study uses MSCI as an empirical rating basis to ensure consistency. The characteristics used by this firm are shown in figure one.

3 Pillars	10 Themes	35 ESG Key Issues	
Environment	Climate Change	Carbon Emissions Product Carbon Footprint	Financing Environmental Impact Climate Change Vulnerability
	Natural Capital	Water Stress Biodiversity & Land Use	Raw Material Sourcing
	Pollution & Waste	Toxic Emissions & Waste Packaging Material & Waste	Electronic Waste
	Environmental Opportunities	Opportunities in Clean Tech Opportunities in Green Building	Opportunities in Renewable Energy
Social	Human Capital	Labor Management Health & Safety	Human Capital Development Supply Chain Labor Standards
	Product Liability	Product Safety & Quality Chemical Safety Financial Product Safety	Privacy & Data Security Responsible Investment Health & Demographic Risk
	Stakeholder Opposition	Controversial Sourcing Community Relations	
	Social Opportunities	Access to Communications Access to Finance	Access to Health Care Opportunities in Nutrition & Health
Governance*	Corporate Governance	Ownership & Control Board	Pay Accounting
	Corporate Behavior	Business Ethics Tax Transparency	

Figure 1- MSCI Key Issue Hierarchy

The Rise in Popularity; Creating Value Through ESG

Through the past two decades there has been a surge of academic research surrounding ESG and its integration in investment strategies. Figure 2 depicts the estimated number of empirical studies on the ESG and corporate financial performance relationship from 1970-2015 (Friede 2015). There is a clear jump in research beginning around the 1995 period, proving how it exploded in popularity especially after its mention in 2006. Overall, investors were realizing implementing ESG analysis into their strategies could create value. A study completed by Henisz (2019) showed one out of every three dollars under professionally managed funds were being invested according to socially responsible principles. Furthermore, between April and June 2020, investment firms which incorporated ESG principles attracted net inflows of $71.1 billion globally, which pushed assets under management within these funds to an all-time high of over $1 trillion (Atkins 2020).

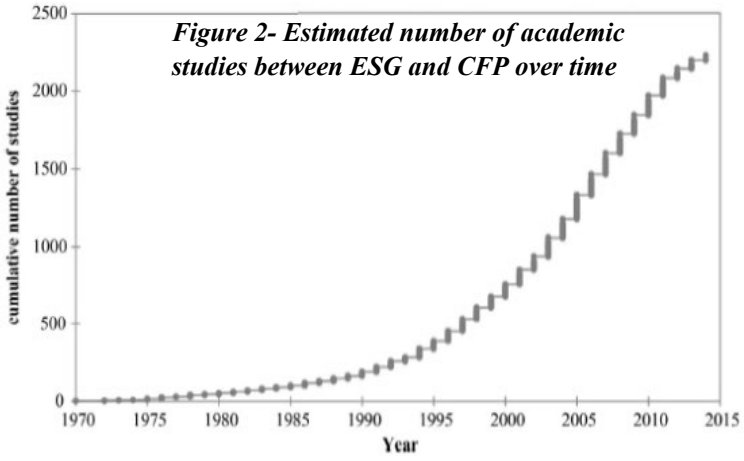


Figure 2- Estimated number of academic studies between ESG and CFP over time

The concept behind this rise in popularity and value creation is the Stakeholder Maximization Theory. Every firm must ask themselves at some point when all is said and done: How do we measure from the past, better or worse? However, this would imply they have some type of performance measure. Value maximization would contend that if the total long run market value of the firm increased, they did well. However, the stakeholder theory would argue that the performance depends on how it considered the interest of stakeholders in the company (Jensen, 2002). With stakeholders being all people or groups associated with the firm like employees, customers communities, or government officials. The realization was that investing with ESG principles can satisfy both theories.

Basically, firms which act accordingly to social standards can create long term value for themselves. By operating with ESG standards in mind, companies are able to please the

employees, customers, and communities (stakeholders) around them. Subsequently, these firms are rewarded with good ESG ratings, and are in a better position to generate top line growth. Government officials are more likely to trust these companies with new projects and approve licenses and resources, allowing firms to tap into new markets and expand into existing ones (Henisz, 2019). Taking all this into account, a firm is able to increase their long term value and please stakeholders with a responsible business practice. By fulfilling the three ESG pillars, they are pleasing the communities and customers around them with morally correct principles. For example, Mckinsey research has found over 70% of customers are willing to pay an additional 5% for a green product if it performs the same as an alternative not as sustainable. Furthermore, they can increase their long term value by having these opportunities to access new resources and expand into new markets. Overall, we are seeing this performance enhancement as a study completed in 2007 revealed nearly a 9% return when purchasing stocks with high ESG ratings and selling those with low ESG ratings (Kempf & Osthoff, 2007).

The Energy Sector and the S&P 500

These studies are commonly inclusive of all sectors in the marketplace, from financials, to technology, to communication. Meaning the portfolios are not restricted to holding only technology companies. However, it is clear each sector are more susceptible to specific ESG pillars over another. For example, a company like Facebook would be more at risk in the social pillar, because they deal with millions, if not billions, of private data points from users. Since this study focuses on energy companies, our portfolios stand more at risk with environmental issues including the scarcity of resources, climate change, pollution, employment, and much more- increasing the need to conform to corporate responsible behavior as both an incentive and a requirement (Stjepcevic & Siksneltye 2017).

The original S&P 500 index was established in 1957 and is widely considered as one of the most used benchmarks for measuring financial performance, as it is constantly updated to include 500 of the leading companies from thriving industries in the economy (Siegal & Schwartz, 2004). Of the twenty largest firms included in these sectors, at the very beginning of the index nine of them were oil companies, all which outperformed the index by 2-3% for over 46 years, showing how important energy companies are for the health of the economy. Today, advancements and trends have shifted the index, and the biggest sectors include information technology and health care (Ross, 2020). As of 2020, the energy sector made up only 2.53% of the S&P 500 index, a steep drop from its original weighting. It consists of nonrenewable energy companies, which mine product like oil, and renewable energy companies. These firms use resources which naturally replenish like sunlight or wind. Since renewable energy companies are relatively new, they have small market capitalization compared to big oil companies, some which have been around for over a century. Therefore, these renewable energy companies have not found their way into the S&P 500 just yet as they do not meet the criteria to be admitted. The breakdown for the energy sector in the S&P 500 can be seen in figure 3.

Sector	Industry	% Of Sector
Energy	Integrated Oil and Gas	50.88%
	Oil & Gas Equipment and Services	8.13%
	Oil & Gas Exploration and Production	20.30%
	Oil & Gas Refining & Marketing	11.51%
	Oil & Gas Storage & Transportation	9.18%

Figure 3- Breakdown of the Energy sector in the S&P 500

Since renewable energy companies are relatively new, they have small market capitalization compared to big oil companies, some which have been around for over a century. Therefore, these renewable energy companies have not found their way into the S&P 500 just yet as they do not meet the criteria to be admitted. The breakdown for the energy sector in the S&P 500 can be seen in figure 3.

DATA

Sample

To begin my analysis, my first steps involved gathering a sample from the massive population of energy companies. I began by obtaining four energy ETF's.

1. XLE- The Energy Sector SPDR Fund
2. XOP- The SPDR Oil & Gas Exploration & Production
3. TAN- The MAC Solar Global Index
4. ICLN- The IShares Global Clean Energy ETF

Both the XLE and XOP include companies that handle nonrenewable energy like coal, natural gas, oil, and nuclear energy. The TAN and ICLN hold companies that handle renewable energy which is not depleted when used, such as wind or solar power. After sorting, my sample was concluded with 78 companies, 43 renewable and 35 nonrenewable. From here, I needed to find reliable financial data for the sample. Using the Center for Research in Security Prices database, I downloaded monthly adjusted returns including dividends, shares outstanding, adjusted closing prices, and tickers for each company from 2018-2021. The four-year window was applied solely because rating agencies change their methods so often, it was important to have the ratings as consistent as possible. At this point I moved on to obtain the returns for my benchmark: the S&P 500. Using the same method I downloaded the monthly returns for the index from 2018-2021.

Measures

My measures included data from an ESG timeseries database (Q1 of 2018) which laid out, for each company in my sample, their most recent ESG scores from Morgan Stanley Capital International. This included a rating on the AAA-CCC scale, their industry adjusted score, weighted score, along with an individual score for each ESG pillar. Figure 4 is an example of a company rated through MSCI.

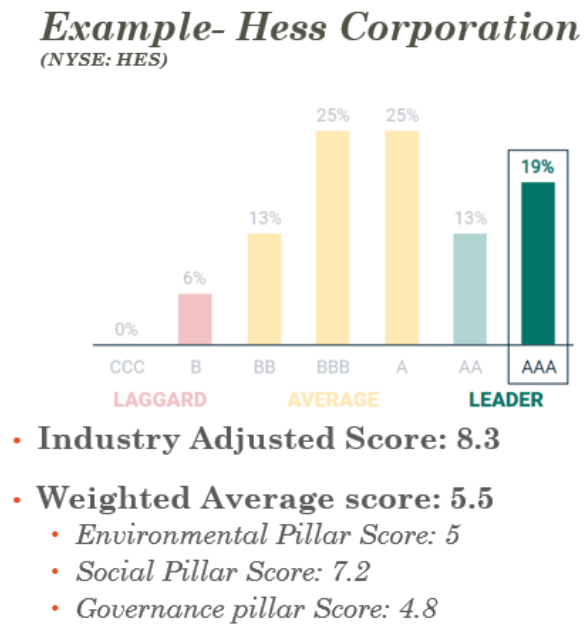


Figure 4-MSCI rating of Hess Corp.

The letter grade shows the highest possible rating for Hess at a triple A. The industry adjusted score is then just a direct translation from that letter grade. The weighted average score is the industry adjusted weighted based on their peers. Finally, the three pillar scores are how the firm fares in each factor.

METHODOLOGY

Once the information was sorted and organized for the 78 companies in our sample, I began to further sort them based on ESG score. I split them between renewable and nonrenewable energy companies, leaving two groups. From here, I found the median industry adjusted score for each group. While the industry adjusted score is a 0-10 rating, the medians both fell below 5. The median score for the nonrenewable group was a 4, and the nonrenewable group was a 4.7. I split each group once again based off if they were above or below this median score. In the end, the sample was split into four portfolios.

1. Renewable energy ESG leaders (21 holdings)
2. Renewable energy ESG laggards (22 holdings)
3. Nonrenewable energy ESG leaders (19 holdings)

4. Nonrenewable ESG laggards (16 holdings)

I then calculated the market capitalization for each company from January of 2018 by multiplying their adjusted closing price by the total outstanding shares. This number is a measure as to how much the market thinks the company is worth at that point in time. After this, I took the sum of the market capitalization from each portfolio, and divided it by each company's market capitalization to get the individual market capitalization weighting for the portfolio. I also took the equal weighting for each portfolio simply by dividing one by the number of holdings in each portfolio. Meaning I had two separate weightings for all four portfolios, which can be seen in the appendix.

Using the monthly returns downloaded from the CRSP I found the monthly holding period return for the portfolio by summing the products of each return by its weighting, giving me two columns of monthly returns. Figure 5 is an example of how the monthly holding period return was calculated for all portfolios.

(Company_A_2018January_Return * Equal_Weight%) + (Company_B_2018January_Return * Equal Weight%)… etc.

(Company_A_2018February_Return * Equal_Weight%) + (Company_B_2018February_Return * Equal Weight%)… etc.

(Company_A_2021December_Return * Equal_Weight%) + (Company_B_2021December_Return * Equal Weight%)… etc.

Figure 5- Example of how holding period returns were calculated

Capital Asset Pricing Model

The Capital Asset Pricing Model (CAPM) is a regression model which describes the relationship between systematic risk and expected return for equities. Essentially, it is a helpful method to price securities fairly taking its risk and time value of money into account. To run the CAPM

model, the risk free rate must be calculated. For this study, the 10 year treasury bill rate was taken monthly through 2018-2021. This is seemingly a "risk free" investment since it is backed by the full faith and credit of the US government. From here, I subtracted the risk-free rate from the monthly holding period returns for each portfolio, to get excess return. This same process was followed with the S&P500 returns, leaving 3 different columns of excess returns shown in figure 6. From here, a regression was ran using one portfolio excess return as our dependent variable, and the S&P excess return as the independent variable, acting as the market return. The regression results display the alpha or excess return on our portfolio compared to the S&P. It also calculates the beta, which is the measure of volatility and risk compared to the market. A beta of 1 shows no risk, while an increase indicates more volatility.

(EW)Portfolio Excess Return	(MW) Portfolio Excess Return	S&P Excess Return
-0.9679%	4.53170%	6.17359%
-6.9399%	-10.64671%	-3.63844%
5.4058%	-0.18249%	-3.13129%
2.7106%	5.16684%	0.91695%
9.3070%	4.14051%	2.42853%
0.7862%	1.47523%	0.12312%
1.8012%	-0.12508%	4.16785%
0.1625%	-0.99449%	3.18966%
-3.0820%	2.99637%	0.13865%
-6.5027%	-8.07846%	-6.49165%
-1.4278%	-1.82339%	1.85241%
-13.6688%	-12.96046%	-9.33656%
14.1155%	10.82848%	8.63509%

Figure 6- A snippet from the leaders, nonrenewable portfolio showing excess returns

RESULTS

In a perfect world, this study would see the greatest returns centered around the renewable leaders in energy portfolio. In this situation, the companies are being morally and socially responsible, stakeholders are satisfied with where their money is going, while in the end both the firm and investors are seeing return. However, different factors must be considered as to why this did not happen in this study.

Empirical Results

The information displayed below in figures 7 and 8 show basic statistics calculated from the holding period returns in each portfolio. The "Avg Return" indicates the average monthly holding period return. Variance is a measure of the dispersion of returns in a portfolio. It is also an indication of correlation between securities in a portfolio, meaning how likely they are to move together.

Nonrenewable			
		Equal Weighted	Mkt Cap Weighted
Leaders	Avg Return	1.95%	0.44%
	Variance	0.9690%	0.914%
Laggards	Avg Return	3.139%	1.019%
	Variance	2.2485%	2.166%

Figure 7- Nonrenewable portfolio statistics

Renewable			
		Equal Weighted	Mkt Cap Weighted
Leaders	Avg Return	1.513%	1.484%
	Variance	1.0010%	1.821%
Laggards	Avg Return	3.315%	2.722%
	Variance	3.1026%	4.162%

Figure 8- Renewable portfolio statistics

The highest average return belongs to the equally weighted renewable laggards portfolio with an average monthly return of 3.315%. The market capitalization weighted portfolios showed a common theme of having a lesser return. In essence, weighing by market capitalization provides the least risk with the greatest amount of potential return by increasing exposure to higher valued firms. On top of this, variance proved to be higher within the laggards of both renewable and nonrenewable.

CAPM Results

Figure 9 shows an example of the regression results after running the Capital Asset Pricing Model on a portfolio. This includes the coefficient for both the intercept and x variable- the

market return. The intercept coefficient indicates the alpha generated, while the variable coefficient is the beta.

Regression Statistics					
Multiple R	0.798097292				
R Square	0.636959288				
Adjusted R Square	0.629067099				
Standard Error	0.061579935				
Observations	48				
ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.306050173	0.30605	80.70755239	1.09673E-11
Residual	46	0.174436066	0.003792		
Total	47	0.48048624			
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*
Intercept	-0.00857123	0.009271246	-0.9245	0.360052704	-0.027233281
X Variable 1	1.598639668	0.177948158	8.983738	1.09673E-11	1.240448604

Figure 9- Equally weighted CAPM regression for the renewable leader's portfolio

The CAPM model also calculates the R-square statistic. This is a representation of the proportion of the variance for our portfolio returns which can be explained by the market return. In this example, the R-squared sits at 63.7%. This indicates the portfolio does not follow the S&P500 performance very well. As if it was closer to 1 or 100%, the portfolio would mimic the index's performance. The R-square sits around this number for all CAPM regressions.

Figures 10 and 11 represent the final CAPM results from each portfolio.

Nonrenewable			
		Equal Weighted	Mkt Cap Weighted
Leaders	Beta	1.5	1.44
	Alpha	-0.2632%	-1.693%
Laggards	Beta	2.03	2.21
	Alpha	0.1251%	-2.259%

Figure 10- Nonrenewable portfolio CAPM results

Renewable			
		Equal Weighted	Mkt Cap Weighted
Leaders	Beta	1.6	2.07
	Alpha	-0.8571%	-1.582%
Laggards	Beta	2.400	2.715
	Alpha	-0.2436%	-1.3024%

Figure 11- Renewable portfolio CAPM results

The only positive excess return, or alpha, was found in the equally weighted, nonrenewable laggards group at 0.1251%. In all other groups it was negative, indicating the market outperformed our portfolios. On top of this, the betas sat relatively high, with the highest being at 2.7. This shows massive volatility, meaning the returns do not move in tandem with the S&P500 at all.

Explanation of Results

The results were not as anticipated but can be explained by a variety of factors. The average return was highest with equally weighted portfolios verse market capitalization weighted. Market capitalization favors companies with high values, generally with little fluctuation. In this sample, there was a large amount of small cap funds, especially within the renewable groups. Meaning the small cap funds performed well over the four-year period, leading to a higher return in equally weighted portfolios. The variances shown in figures 7 and 8 were generally higher in the laggards group in both renewable and nonrenewable energy. Basically, this shows these stocks with lower ESG ratings have more dispersion than higher rated firms. It also indicates they have little correlation with each other, meaning predictability of return comovement is much lower.

The CAPM regression results showed the only positive excess return compared to the S&P 500 in the nonrenewable laggard's portfolio. While it goes against the stakeholder maximization theory, it can be attributed simply to sheer size and history in large oil firms. An article published by Desilver (2020) described how fossil fuels continue to dominate the US economy. In 2018, fossil fuels led the US economy by feeding about 80% of the nation's energy demand, versus solar and

Environmental_Pillar_Score	Social_Pillar_Score	Governance_Pillar_Score
2.9	6.5	6.1
4.4	6.5	6.2
4.3	1.4	4.7
6.7	3.7	5.8
1.9	3.6	5.5
1.9	3.4	4.9
6.7	3.1	4.5
2.3	6.8	5.4
2.9	5.6	5.4
2	3.3	4.9
3	2.7	3.9
1.7	3.3	5
4.9	7	1.9
4.9	0.6	6.3
2.5	6.5	5.5
1.2	4.4	4.3

Figure 12- Nonrenewable laggard energy firms ESG pillar scores

wind energy supplying barely 4% (Desilver 2020). We can see that although renewable companies are on the rise, coal, oil, and natural gas were still the go to energy sources through the 2018-2021 period observed. Additionally, with regards to ESG, we saw poor Environmental pillar scores within this portfolio, with higher social and governance scores shown in figure 12. The reason as to why it was able to generate return in this sample was due to the small cap firms. Since it was equally weighted, the volatile small cap firms performed well through the four years which contributed a lot to the excess return.

Within the renewable energy group, the results showed a lot of volatility. Since many of the renewable firms are small caps, standard deviations proved to be rather high for each firm with some exceeding 50%. It is a representative of the size and performance comparison of renewable and nonrenewable firms.

CONCLUSION

Overall, there was no clear outperformer through this process. While there was only one portfolio producing a positive alpha, it was not high enough to justify that the sample can speak for the population. On top of this, the R-square was only at .45, meaning not even 50% of the returns on the nonrenewable laggards portfolio could be explained by the market returns.

This project aimed to use solely ESG scores as a characteristic to construct portfolios, centered in the energy industry. The results show a great deal of volatility, with little excess return compared to the S&P 500. While more research should be done, it shows that in the energy sector ESG should only be a helping factor in investment strategies, not the end all be all. We are seeing a consistent, dominant performance in oil and gas companies compared to renewable energy firms, a trend which will take a long time to dissipate.

The results additionally make a case for the agency cause theory, which is the complete opposite of the stakeholder theory. In essence, it argues firms should not spend resources on things such as environmental, social, and governance activities as it will decrease shareholder value (Peng 2020). It believes that it will take away from more important aspects which will generate profits. However, it neglects sustainability and responsible corporate behavior.

The biggest issue with this project is the time period, as it goes through the Covid-19 pandemic. During this time, the stock market, experienced a great deal of fluctuation, interest rates plummeted to near zero, and hundreds of firms failed. This was not a “normal” market. Meaning the results displayed through these portfolios would be difficult to translate to an efficient market.

Explainability of Predictors in Asset Pricing Using Machine Learning and LIME

Bhaskar Goswami[a], Ajim Uddin[a]

[a]*Martin Tuchman School of Management, New Jersey Institute of Technology, USA*

Abstract

A vast majority of asset pricing literature is focused on understanding the capacity and explainability of asset pricing models and their respective predictors. In this paper, we take a new take on this aspect using explainable AI. We analyze over 200 asset pricing predictors previously identified in the literature to understand their importance in various asset pricing model settings. Our findings suggest that Bagging and Boosting-based models have an edge in providing generalized predictions across various dependent variables. We also understand that some of the features commonly used as good predictors are interoperable. Using an advanced machine learning algorithm - LIME, we also find that momentum and trading-based features have higher predictive powers in estimating asset returns. The finding is consistent across multiple models and after controlling for characteristics-based predictors.

Keywords: Asset Pricing, Machine Learning, Explainable AI, Model Predictability, FinTech

1. Introduction

The use of Machine Learning and Data Science in the field of Computer Science and Engineering has been prevalent for a long time. In recent years,

Preprint submitted to NEDSI 2023 Annual Conference

the use of this technology has been growing in popularity in various intradisciplinary fields due to collaboration and sharing of interest and knowledge.

There has been a major advancement in the field of finance and time series forecasting of returns and predicting the prices of various assets. According to Weigand (2019), Machine Learning can be defined as a collection of complex models, combined with methods of regularization and other methods to prevent over-fitting accompanied by algorithms to search for the optimal specification of these algorithms commonly known as hyper-parameter tuning.

The use of model interpretability tools has recently been a popular topic among researchers. People earlier were more concerned about how to get accurate predictions. The industry and people using these technologies currently need to know how these algorithms work and how the predictions can be used effectively in practice. This is answered by various frameworks that are being developed to enhance model explainability and drive the adoption of these technologies in every domain. The most common ones being used are Shapely and LIME. We will be using LIME along with other Machine Learning models, algorithms, and techniques in our paper to get an idea of which features are the key to predicting Empirical Asset Pricing and Returns for various companies across a timeline from 2000 to 2022. This provides us with a robust viewpoint on model interpretability and increases the transparency in the models that are otherwise considered a black box in today's world.

We needed to investigate the extant literature and get an elaborate idea of the domain and the various methodologies that are applied in the field of Asset Pricing with Machine Learning. Scopus, a widely used database for

publications, was used to systematically filter out documents related to the research topic.

To be true to our current circumstances, we used only those papers that were already published in reputed journals and were written in English for our better understanding. Interestingly, most of the papers were from developed countries with huge advancements in Financial Technology and Data Science in general. In Figure 1 we can see the distribution of journal articles from across the globe. The oldest article that we reviewed was Liu and Fu (2016). As the industry of Finance and Technology is evolving at a very rapid rate, we did not find it necessary to look at publications that were more than half a decade old. This gave us an idea of the current situation in the industry at large.

We finally reviewed 75 articles; Table 1 shows the relevant statistics for them. This was used to get a brief idea about the domain, and to find and examine gaps in extant research. Formulating effective questions to deal with those research gaps has become particularly important in recent times when people need to know the answer to "How?" rather than "What?".

The various sources provided information about the use of Machine Learning and Deep Learning in Empirical Asset Pricing Gu et al. (2020). It ranged from simple linear models such as Linear Regression to models that were non-parametric such as Decision Trees, Random Forests, and other Ensemble techniques. It also showed the usefulness of Multilayered Neural networks and their effectiveness in the predictability of asset pricing or stock price, in general, Gu et al. (2020). Many publications talked about the various predictors

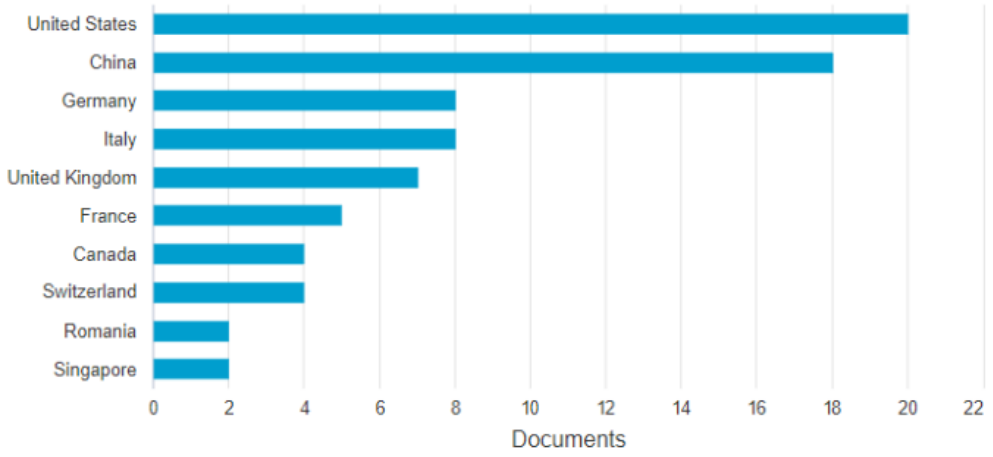


Figure 1: The vertical axis represents the various countries involved in research in this domain. More than a quarter of the articles were from or about the United States of America.

Table 1: Descriptive statistics of articles included in the Literature Review

Description	Metrics
Time span	2016:2022
Sources	62
Documents	76
Average Years from Publication	1.55
Average citations per Document	10.25
Indexed Keywords	647
Authors Keywords	390
Authors	217
Single Author Documents	13
Authors per Document	2.855

and which one of them is important for asset pricing Huang et al. (2022), D'Amato et al. (2021).

There were articles that investigated the curation of portfolios using deep learning and neural networks. Additionally, some papers ran analyses on unstructured data such as news feeds Meyer et al. (2017), Atkins et al. (2018),

and social media Houlihan and Creamer (2021), Houlihan and Creamer (2017) to predict the volatility of returns. They employed sentiment analysis for the same. In asset pricing researchers also tried to predict the pricing of different instruments such as bitcoin Jaquart et al. (2021), Jaquart et al. (2020) other cryptocurrencies Wang (2021), options Houlihan and Creamer (2017), Goudenège et al. (2019), Goudenege et al. (2020), Eckstein et al. (2021), securities Yi (2022), Baker et al. (2020), Deku et al. (2021), and appraising real estateWeigand (2019) in various markets such as China Jiang et al. (2022), Zhao et al. (2022), Liu and Fu (2016), Europe Drobetz and Otto (2021), and the US.

The various research questions that were asked while doing the literature review and thinking about the research gaps were: 1) How big of a part does Data play in the field? 2) What are the various techniques and methods applied to the datasets to get an accurate forecast? as seen in Li and Mei (2020) 3) What are the various target variables that researchers aim to predict and what is their impact on business and the industry at large? 4) How can we increase the transparency and interpretability of these models?

The research gap that we intend to address through this paper is the introduction of explainability and transparency in the feature and model selection process of Empirical Asset Pricing through Machine Learning. We aim to provide a list of open-source features that have high predictive power when coupled with specific machine-learning algorithms. In addition to that, we intend on providing cogent and coherent explanations for the selection of models and features along with the comparison of performance between various algo-

rithms to get a holistic overview of the field of Machine Learning in Asset Pricing and Returns. Throughout the paper, we try to figure out what should be the ideal set of attributes or features to be considered while predicting the return or excess return of stocks of various companies across a wide time horizon and look at the bigger picture. In this paper, we intend to provide suggestions in the form of lists of predictors that are potent for various use cases or problem statements in Empirical Asset Pricing.

2. Methodology

2.1. Data Collection

In academia, the field of asset pricing and finance goes through a roller coaster of teamwork and close competition Chen and Zimmermann (2021). The secondary data that we have considered in this research project is the Open-Source Cross-Sectional Asset Pricing data made available to the entire world by Dr. Andrew Y. Chen and Dr. Tom Zimmermann. The data has more than 200 predictor variables (that are mentioned in the extant literature as proven predictors) and 130 placebo variables (that may not have much predictive power). The features present in the dataset have been calculated on a month-on-month basis. The target variables that we have considered for our study are the Return, the Excess Return from the Market Model, the Excess Return from the Cahart Four-Factor Model, and the Excess Return from Fama and French Three-Factor Model.

2.2. Models Used

We plan to employ machine learning models to get an idea of which of these predictors are valuable across a varied set of models. We considered simple models such as Linear Regression, Decision Tree Regressor, some Ensemble models such as Light Gradient Boosting Machines (LGBM) and Extreme Gradient Boosting (XGBoost) Machine Learning Algorithms, and an assortment of Neural Network architectures with 4 and 5 hidden layers. As per the pyramid rule, the 4 hidden layered neural network contains 32, 16, 8, and 4 nodes whereas the 5 hidden layered neural network has 32, 16, 8, 4, and 2 nodes in their hidden layers according to Masters (1993). The output layer is one single node providing the prediction for the dependent variables.

We intend on using LIME and underlying Shapely values Aas et al. (2021) to better understand the models, their predictions, and the importance of each predictor variable separately. In this paper, we concentrate on finding the best models for predicting the different types of returns that are prevalent in the domain using the models and features that are used in the empirical asset pricing literature. We intend to provide an approach for model explainability of the features by providing a measure of their contribution to the variation in the predictions. We do not intend to do any kind of hyperparameter tuning or tweaking neural network architectures. What we hope to achieve is the lower bound of the performance that these models can provide without tuning them or using any reasonable validation pipeline.

2.3. Evaluation Framework

For evaluating and comparing the results of various models we put in place a framework that has been widely used in the extant literature and is a frequent practice in Machine Learning and Data Science at large Gu et al. (2020). We plan to split the data set into a training sub-sample and a test sub-sample. The split will be done based on the time stamp that includes the month and the year in our case. The training sample is used to train the various models and the test set is used to simulate the OOS (Out of Sample) set for each model Gu et al. (2020). The metric widely used in the publications for this domain is the Out of Sample R2 score (3) adapted from Gu et al. (2020). In addition to that, we also looked at other metrics like MAE (Mean Absolute Error) (1) and MSE (Mean Square Error) (2) to have a well-rounded idea of the model performance. The value of R2 OOS ranges from 0 to 1, where 1 is the best model and 0 is the worst. The values for MAE and MSE are negatively correlated to the performance of a machine learning predictive model.

$$MAE = \sum_{i=1}^{N} |Actual_i - Predicted_i| \tag{1}$$

$$MSE = \sum_{i=1}^{N} (Actual_i - Predicted_i)^2 \tag{2}$$

$$R^2_{oos} = 1 - \frac{\sum_{i=1}^{N} (Actual_i - Predicted_i)^2}{\sum_{i=1}^{N} Actual_i^2} \tag{3}$$

3. Data Processing, Model Building, and Explainability

The secondary data consisted of more than 200 predictors that have been known to work on the use case of Empirical Asset Pricing throughout the literature. As a dependent variable, we look into Actual Returns and three kinds of Excess Returns (the Excess Return from the Market Model, the Excess Return from the Cahart Four-Factor Model, and the Excess Return from the Fama and French Three-Factor Model).

There were many predictors in the data that had missing values. Handling missing values is one of the major concerns of every predictive problem statement. We considered the data after the year 2000 to eliminate some of the noise and the NULL values from our dataset. We went from having more than 34,000 unique companies to around 15,000 unique companies in the dataset. After that, we tried to apply a forward-filling technique based on Company ID (or "permno" as per our secondary dataset). Here, we did null imputation for a particular company in a particular year with data that was presented earlier for that company in the dataset.

In addition to that, we performed a NULL analysis filtering out companies that have more than 50 percent of their attributes completely NULL and then removed them from our dataset as they will not be useful to our predictions and do not provide any variance to the models. In addition to that, we removed all individual rows that have more than or equal to 70% NULL values. We were left with a little over 10,000 company IDs in a total dataset of 1,000,000+ rows and 166 predictors for the model building, prediction, and evaluation process.

To look at the performance of models, we divided the data into train and

test based on date. Any record before 2018 was put in the train set and the rest in the test set for evaluating the model. As we do not have any hyperparameter tuning in the scope of the project, we do not need the validation set or any other validation pipeline to tune the model parameters.

We used LIME (Local Interpretable Model-Agnostic Explanations) to help with model interpretability. For all the models, we can see the predictions for a random row in the test set. As the feature importance in LIME is in the unit of the predictions, we can say accurately which feature provided how much of the variance from the baseline for each prediction made by our models.

4. Results

The following section contains the results that we got by evaluating different models across our four different dependent variables considering all the 166 features or predictors that were part of our independent variable set after the initial preprocessing.

We can see the results on the training data as well as the testing data. It is evident from the results that the boosting algorithms provide better results but are prone to over-fitting. However, linear models and other simpler models such as Linear Regression have an under-fitting problem where they cannot catch the intricate, inherent, non-linear properties of the data. Decision Tree Regressors, on the other hand, when not restricted with hyperparameters overfit the data by memorizing the entire train set data intricacies and this causes the results on the train set to be extremely high and perform poorly on the test data. This is a limitation for our research where we have not looked into

the hyperparameter tuning aspect of the model building pipeline which is one of the most important, computationally heavy, and time-consuming tasks for any problem statement. We can see the results for various models for various independent variables in Table 2

For the Neural networks, we can see a good generalization of the prediction where in the Return from Fama and French Three-Factor Model in Panel D of Table 2 the metrics for both the train and the test set are close to one another. This suggests that with proper hyperparameter tuning, more epochs (we only iterated through 100 epochs), and adding other techniques like dropout and early stopping we can get better results with a multi-layered neural network.

A very similar view can be presented when we look at the results for the various predictions with the help of Gradient Boosting models. They are generalized and provide better results out of the box without extensive hyperparameter tuning. If we optimize the hyperparameters we can get better results with models that are less complicated than that of the black box of Neural Networks. However, we need to keep in mind that the Neural networks that we used were very shallow, i.e., only 4-layered and 5-layered. We can use more hidden layers and more complex neural networks to capture more non-linear intricacies of the data if present and provide better overall performance.

The top ten features while predicting each dependent variable by each model can be seen in Table 3.

From the results, it is evident that the models needs to go through hyperparameter tuning to get rid of the overfitting condition as seen in the case of the decision tree. Trimming and Pruning of the decision tree are needed to

Table 2: Results Summary

	Train			*Test*		
Models	MSE	MAE	R^2	MSE	MAE	R^2
Panel A: Return						
Linear Regression	184.625	8.162	0.427	398.013	9.648	0.128
Decision Tree	0	0	1.0	366.277	9.523	0.219
LightGBM	67.651	5.129	0.790	257.514	7.261	0.435
XGBoost	45.609	4.591	0.858	266.479	7.175	0.416
ModelNN4	150.80	7.404	0.53	771.36	9.631	−0.689
ModelNN5	323.138	10.766	−0.001	456.507	11.304	−6.159
Panel B: Excess Return from Cahart Four-Factor Model						
Linear Regression	208.420	8.636	0.327	389.033	9.784	0.109
Decision Tree	0	0	1.0	468.309	11.716	−0.054
LightGBM	110.258	6.654	0.643	299.434	8.508	0.315
XGBoost	80.925	6.083	0.738	303.541	8.499	0.305
ModelNN4	186.68	8.666	0.397	428.53	10.910	0.019
ModelNN5	202.34	9.117	0.346	470.19	11.436	−0.075
Panel C: Excess Return from Market Model						
Linear Regression	183.424	8.088	0.378	372.796	9.537	0.111
Decision Tree	0	0	1.0	395.955	10.555	0.048
LightGBM	78.865	5.721	0.732	279.789	8.133	0.332
XGBoost	55.058	5.129	0.813	284.561	8.034	0.321
ModelNN4	155.374	8.052	0.473	402.637	10.888	0.040
ModelNN5	296.633	10.252	−0.005	423.456	10.710	−0.009
Panel D: Excess Return from Fama and French Three-Factor Model						
Linear Regression	194.842	8.370	0.343	373.992	9.571	0.115
Decision Tree	0	0	1.0	472.687	11.193	−0.141
LightGBM	94.748	6.283	0.680	282.357	8.267	0.332
XGBoost	69.342	5.731	0.766	289.062	8.274	0.316
ModelNN4	174.847	8.587	0.411	405.209	10.999	0.041
ModelNN5	299.499	10.387	−0.008	425.662	10.611	−0.007

Table 3: Important features for each Independent Variable and Model Combination

Model \ Y	Return	Excess Return from Market Model	Excess Return from Fama and French Three-Factor Model	Excess Return from Cahart Four-Factor Model
Decision Tree	High52 Mom12m MaxRet ReturnSkew Mom6m IndRetBig IdioRisk Beta Mom12mOffSeason IdioVolAHT	High52 Mom12m MaxRet ReturnSkew Mom6m IdioRisk Mom12mOffSeason IdioVolAHT IdioVol3F MomSeasonShort	High52 Mom12m MaxRet ReturnSkew Mom6m IdioRisk Mom12mOffSeason IndRetBig MomSeasonShort IdioVol3F	High52 Mom12m MaxRet ReturnSkew Mom6m Mom12mOffSeason IdioRisk IndRetBig IdioVol3F Beta
LightGBM	High52 Mom12m Mom6m MaxRet Mom12mOffSeason IdioVolAHT IdioRisk IndRetBig IntMom ReturnSkew	High52 Mom12m Mom6m IdioVolAHT MaxRet Mom12mOffSeason IdioRisk Beta IdioVol3F IndRetBig	High52 Mom12m Mom6m MaxRet IdioVolAHT Mom12mOffSeason IdioRisk IndRetBig Beta IdioVol3F	High52 Mom12m Mom6m IdioVolAHT Mom12mOffSeason MaxRet IndRetBig IdioRisk Beta IdioVol3F
XGBoost	High52 Mom12m ReturnSkew IndRetBig MaxRet Beta Mom6m IdioVol3F IdioRisk Mom12mOffSeason	High52 Mom12m ReturnSkew MaxRet IdioVol3F Mom6m IdioRisk InvestPPEInv ShareVol Mom12mOffSeason	High52 ReturnSkew Mom12m MaxRet IdioVol3F IdioRisk Mom6m Mom12mOffSeason Beta ReturnSkew3F	High52 ReturnSkew Mom12m MaxRet IdioVol3F IdioRisk Mom6m Mom12mOffSeason ReturnSkew3F ShareVol

prevent it from overfitting. It is also clear that the boosting algorithms provide more of a generalized model for the prediction of the Returns and Excess Returns.

Regarding the feature selection process, we can see many common variables that appear in the list of top 10 features that are important for the models to make accurate predictions. For instance, the variable “High 52” and the “Mom12m” appear in the top three of every model for every depen-

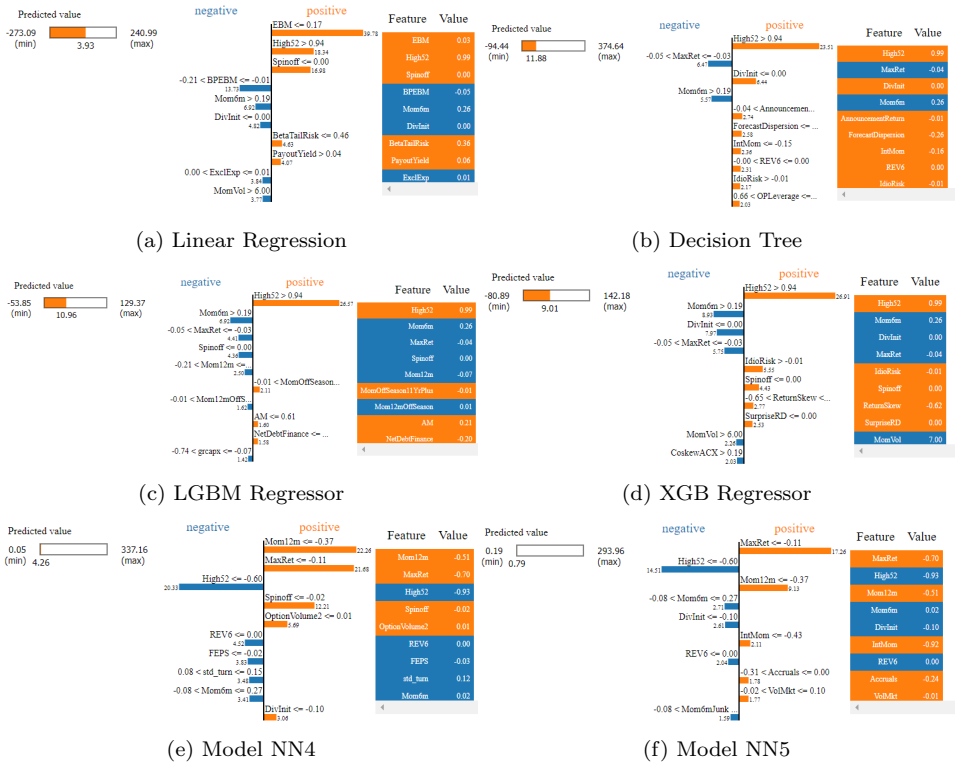


Figure 2: LIME used to make the prediction of Returns by various Machine Learning Models Interpretable. Each subplot shows the prediction, the top 10 features, and their contribution to the variance that results in the prediction of a random sample in the test set.

dent variable. This provides us with an important insight into the selection of important predictors in the domain of Asset Pricing. It helps practitioners to realize that momentum and trading-based features hold more importance in the empirical asset pricing and prediction of the return of various assets.

The use of LIME for the purpose of model explainability can be seen in Figure 2a we can see the Linear Regression model predictions for a random instance from the test set that the model has not been trained on and shows the model explainability via the feature importance of each feature in the unit

of the prediction itself. For instance, the attribute named High52 has an 18.34 contribution and Mom6m has a negative 6.92 contribution to the resulting variance in prediction from the baseline or mean value of the data.

In the corresponding images (Figure 2b, Figure 2c, Figure 2d, Figure 2e, Figure 2f) we can see LIME explaining similar predictions on the test set for the other models that we have considered including the complex multi-layered neural networks. We have presented only the ones for RET (Return), and this can be replicated for all the dependent variables for the dataset that we have considered with similar explainability features.

5. Practical and Managerial Implications

This line of research has great practical and managerial implications. For the hedge fund manager, it is easier to understand the trends in the various key features of the asset pricing domain and make quick decisions based on algorithms that are well-established in the field. For the average person, trying to understand the various attributes that play a key role in stock price and returns and by what magnitude is key to making an informed decision. The best thing about frameworks like Shapley and LIME is that the results show the variable or predictor's importance on the scale of the predicted value so that there is a clear explanation of every variance that the model produces as predictions as the additive measure of all the variances introduced by all the features involved.

In the current times when technological advancements are having a counter-intuitive reaction to the world at large, where people are skeptical of the

question "How?" The adoption and explainability of technology in various fields is an overarching issue. We are trying to tackle this issue through this research project. This will help with the adoption of Machine Learning and Data Science in the field of Financial Technology where it will not only play as a plug-and-play tool or a black box but provide information into the prediction process and the intuition or insights behind those predictions to the common person or the trained professional taking a deeper dive into the models and their predictions for making a data-driven and informed decision.

6. Conclusion and Future Scope

In the current scenario, the use of machine learning, predominantly deep learning techniques in the domain of finance has been a very engaging challenge for many researchers and scholars. There have been innovations and novel ideas in using varied kinds of structured and unstructured data to model the return from various assets and for pricing of those assets. There is a need to use news and other social media data to gauge the importance of people and their sentiments on the stock market and various other assets.

With the advancement of technology, there has been a need for more explainability and transparency of these algorithms and various predictors used by various scholars and practitioners. This is an area of research that we feel is being neglected and needs proper attention from both academia and industry.

In the transition from "What?" to "How?" and "Why?" questions, people are now more concerned about the societal impact of a particular line of research. The rise of explainable AI gives the common person an avenue to

understand complex algorithms and make an informed decision that affects his/her life considerably. The transparency of overly complex algorithms and their adoption into empirical asset pricing where models are not treated as a black box anymore has become imminent.

References

Aas, K., Jullum, M., Løland, A., 2021. Explaining individual predictions when features are dependent: More accurate approximations to shapley values. Artificial Intelligence 298, 103502.

Atkins, A., Niranjan, M., Gerding, E., 2018. Financial news predicts stock market volatility better than close price. The Journal of Finance and Data Science 4, 120–137.

Baker, A., Gelbach, J.B., et al., 2020. Machine learning and predicted returns for event studies in securities litigation. Journal of Law, Finance, and Accounting 5, 231–272.

Chen, A.Y., Zimmermann, T., 2021. Open source cross-sectional asset pricing. Critical Finance Review, Forthcoming .

Deku, S.Y., Kara, A., Semeyutin, A., 2021. The predictive strength of mbs yield spreads during asset bubbles. Review of Quantitative Finance and Accounting 56, 111–142.

Drobetz, W., Otto, T., 2021. Empirical asset pricing via machine learning: evidence from the european stock market. Journal of Asset Management 22, 507–538.

D'Amato, V., D'Ecclesia, R., Levantesi, S., 2021. Fundamental ratios as predictors of esg scores: a machine learning approach. Decisions in Economics and Finance 44, 1087–1110.

Eckstein, S., Guo, G., Lim, T., Obłój, J., 2021. Robust pricing and hedging of options on multiple assets and its numerics. SIAM Journal on Financial Mathematics 12, 158–188.

Goudenège, L., Molent, A., Zanette, A., 2019. Variance reduction applied to machine learning for pricing bermudan/american options in high dimension. arXiv preprint arXiv:1903.11275 .

Goudenege, L., Molent, A., Zanette, A., 2020. Machine learning for pricing american options in high-dimensional markovian and non-markovian models. Quantitative Finance 20, 573–591.

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. The Review of Financial Studies 33, 2223–2273.

Houlihan, P., Creamer, G.G., 2017. Can sentiment analysis and options volume anticipate future returns? Computational Economics 50, 669–685.

Houlihan, P., Creamer, G.G., 2021. Leveraging social media to predict continuation and reversal in asset prices. Computational Economics 57, 433–453.

Huang, J., Xing, R., Li, Q., 2022. Asset pricing via deep graph learning to incorporate heterogeneous predictors. International Journal of Intelligent Systems 37, 8462–8489.

Jaquart, P., Dann, D., Martin, C., 2020. Machine learning for bitcoin pricing-a structured literature review., in: Wirtschaftsinformatik (Zentrale Tracks), pp. 174–188.

Jaquart, P., Dann, D., Weinhardt, C., 2021. Short-term bitcoin market prediction via machine learning. The journal of finance and data science 7, 45–66.

Jiang, Z., Zhang, L., Zhang, L., Wen, B., 2022. Investor sentiment and machine learning: Predicting the price of china's crude oil futures market. Energy 247, 123471.

Li, W., Mei, F., 2020. Asset returns in deep learning methods: An empirical analysis on sse 50 and csi 300. Research in International Business and Finance 54, 101291.

Liu, X., Fu, H., 2016. Volatility forecasting for interbank offered rate using grey extreme learning machine: The case of china. Chaos, Solitons & Fractals 89, 249–254.

Masters, T., 1993. Practical neural network recipes in C++. Academic Press Professional, Inc.

Meyer, B., Bikdash, M., Dai, X., 2017. Fine-grained financial news sentiment analysis, in: SoutheastCon 2017, IEEE. pp. 1–8.

Wang, Q., 2021. Cryptocurrencies asset pricing via machine learning. International Journal of Data Science and Analytics 12, 175–183.

Weigand, A., 2019. Machine learning in empirical asset pricing. Financial Markets and Portfolio Management 33, 93–104.

Yi, K., 2022. The multifactor quantitative investment model based on association rule mining and machine learning. Wireless Communications and Mobile Computing 2022.

Zhao, W., Yang, H., Zhou, H., 2022. Linguistic specificity and stock price synchronicity. China Journal of Accounting Research 15, 100219.

Exposing investor sentiment in stock returns

Abstract

This paper examines the relationship between investor sentiment and contemporaneous stocks returns. We find that sentiment shock had a higher explanatory power on stock returns than sentiment itself, and sentiment shock beta exhibits a much higher statistical significance than sentiment beta. The explanatory power of sentiment and sentiment shock varies in different market states. Compared with sentiment level, sentiment shock has a stronger linkage to the market factors, suggesting that sentiment shock is more responsive to stock returns. We conclude that sentiment shock is a better sentiment factor for stock returns.

Key words: Sentiment, sentiment shock, stock returns, sentiment beta, market factors.

Introduction

From the tech bubble burst to the 2008 financial crisis to the 2020 market crash caused by Covid-19, investor sentiment demonstrated its power in asset returns repeatedly. As the capital market history shows, investors get irrational rather quickly. The relationship between investor sentiment and stocks returns falls into behavioral finance, which is aimed at figuring out the influence of psychology on the behavior of investors and asset prices. Behavioral finance started partly in response to the efficient market hypothesis (EMH), a popular theory that assumes asset prices reflect all information and investors are rational. In real life, however, humans aren't always rational, and the decisions they make might be flawed. After the Prospect Theory was introduced by Kahneman and Tversky (1979), numerous studies have been published in behavioral finance.

De Bondt and Thaler (1985) found that most investors tend to overreact to unexpected and dramatic news events, and such behavior affects stock prices. Solt and Statman (1988) found that

the Bearish Sentiment Index is useless as an indicator of forthcoming stock price changes. [1] De Bondt (1993) also found that the AAII sentiment index could forecast future stock returns.[2] Fleming, Ostdiek, and Whaley (1995) concluded that the CBOE volatility index VIX embeds market expectations. Grinblatt, Titman, and Wermers (1995) found that about 77% of mutual funds were "momentum investors." Clarke and Statman (1998) report no statistically significant relationship between the Investors Intelligence survey and subsequent stock returns. Otoo (1999) found a strong contemporaneous relationship between the changes of consumer confidence index and the stock returns. Goetzmann, Massa, and Rouwenhorst (2000) suggested that investor behavior provides significant incremental explanatory power in the cross-sectional regressions on daily asset returns. Brown and Cliff (2004) find that sentiment levels and changes are strongly correlated with contemporaneous market returns, their tests show that sentiment has little predictive power for near-term future stock returns.

Wang et al. (2006) find that most of the sentiment measures are caused by returns and volatility rather than vice versa, and they find that lagged returns cause volatility. Lemmon and Portniaguina (2006) concluded that consumer confidence can forecast small stocks' returns of and those with low institutional ownership. Baker and Wurgler (2006) found that the cross-section of future stock returns is conditional on beginning-of-period proxies for sentiment. Qiu and Welch (2006) did not detect any relationship between the discount of the closed-end fund and the UBS/Gallup sentiment survey. Cipollini and Manzini (2007) showed that the CBOE volatility index VIX is an important driver of the S&P 500's future returns. Beaumont et al. (2008) found that sentiment has a significant and asymmetric impact on market volatility. Feldman (2010) calculated a sentiment measure called the perceived loss index that can determine perceived risk for different categories of equities. Statman (2011) found an evident relationship among investor sentiment, stock characteristics, and returns. Amihud and Goyenko (2013) found that lower R^2 indicates higher selectivity and significantly predicts better mutual fund performance. Sibley, Wang, Xing and Zhang (2016) showed that the Baker-Wurgler Index is strongly correlated with business cycle variables, suggesting that the power of this sentiment index is mainly driven by the risk/business cycle component. Wulfmeyer (2016) finds that mutual fund managers display a

[1] This index is calculated as the ratio of the number of investment advisers who are bearish to the total number of advisers who are either bearish or bullish.
[2] AAII represents the American Association of Individual Investors.

stronger disposition-driven behavior when stocks are more difficult to value. Similarly, Chu, Du, and Tu (2017) argued that the Baker-Wurgler Index contains too much fundamental factors. Feldman and Liu (2017) concluded that correlations of sentiment can significantly forecast future one-year stock market return correlations and have more substantial predictive power for future return correlations during bear periods. Feng et al. (2017) found that technical indicators perform better during periods of high sentiment than during periods of low sentiment, especially for small stocks. Dong and Doukas (2018) found that the skill of institutional investors matters most during high sentiment periods when market signals are noisy. Bu and Forrest (2020) find that the AAII sentiment indexes can well explain concurrent fund performance by itself, whereas the BW index does not have such an ability. Bu (2021) found that the AAII indexes have significant explanatory power on contemporaneous stock returns, and investor sentiment is mainly driven by abnormal emotions, either bullish or bearish.

Numerous sentiment measured have been created by scholars and practitioners, such as the AAII indexes, the University of Michigan Consumer Sentiment Index, the Bake and Wurgler (2006) index, the US consumer confidence index, Barron's sentiment index, CNN's *Fear and Greed* indexes, the NYSE high/low indicator, the NYSE 200-day moving average, the Commitment of Traders report, and more. These sentiment measures can be categorized into direct measures and indirect measures. Direct sentiment measures are based on information gained through surveys, which seek information from individuals regarding their feelings about the stock market and economic conditions. Indirect sentiment measures are calculated based on financial and economic variables.

The purpose of this study is to find a sentiment measure that better explains contemporaneous stock returns. Specifically, we want to test whether sentiment level or sentiment shock matters more for stock returns. Sentiment shock refers to the change in investor sentiment from the last period. We have noticed that stock market could remain stable at high-sentiment levels, either bullish or bearish, whereas the market could be very sensitive if investor sentiment level changed. Considering this phenomenon, we want to find out whether sentiment shock is a better sentiment measure than sentiment level. Both direct and indirect sentiment measures are available to study this topic. However, the problem with the indirect sentiment measures is that they are derived from market variables, and so they capture the sentiments that are already

exhibited by economic indicators. Thus, we decide to use an indirect sentiment measure to better catch the emotional part of investor sentiment and make a comparison between sentiment and sentiment shock. As mentioned earlier, there are many sentiment measures, either direct or indirect measures. After testing a variety of indirect sentiment measures, we find that the American Association of Individual Investors (AAII) index is a good match for the study.

The rest of the paper is organized as follows: We first present the theory that explains the difference between investor sentiment and sentiment shock. Then we introduce the data and report our empirical findings. The conclusion summarizes the whole paper.

Data

we use the CRSP stock database to construct our stock sample. The sample includes all the U.S. stocks that existed in the sample period. The sample period spans from January 2009 to December 2020, covering a period of 12 years. We include only those stocks with a complete monthly return record in the entire sample period. The stock sample has 3,340 stocks in total.

The American Association of Individual Investors (AAII) index is a typical survey-based sentiment measure. The AAII sentiment index comprises three indexes, including the bullish, bearish, and neutral indexes. This allows us to measure sentiment and sentiment shock from three dimensions. The AAII surveys investor sentiment weekly, measuring the percentage of investors who are bullish, bearish, and neutral on the stock market for the next six months. The three sentiment indexes add up to 100%. In this study, we use the AAII indexes to proxy for direct sentiment measures. The AAII sentiment indexes are downloaded from the website of the American Association of Individual Investors.

Empirical Findings

a. Explanatory power of investor sentiment and sentiment shock

In this section, we compare the explanatory power of investor sentiment and sentiment shock on contemporaneous stock returns. If sentiment shock is more relevant to stocks returns, it must exhibit a higher explanatory power than investor sentiment itself. Since AAII indexes include

three sentiments, we can compare the explanatory power of bullish, bearish, and neutral sentiments on stock returns. The explanatory power is proxied by the adjusted R-Squared values of the following model:

$$R_{i,t} = \alpha_i + \beta_i * S_t + \varepsilon_i \quad (1)$$

Where $R_{i,t}$ is the return of stock i in month t. S_t stands for a sentiment or sentiment shock of the three AAII sentiment indexes. Table 1 report the results.

[Insert Table 1 Here]

As Panel A in Table 1 shows, during the whole sample period, the bullish sentiment has the highest adjusted R-squared with a mean of 2.96% and a median of 2.31%. The mean and median of the bearish index are 1.59% and 0.75%, respectively. The explanatory power of the neutral index is very little with a mean of 0.25% and a median of -0.12%. We noticed that the mean value of the three sentiments is much higher than the median, indicating that the distributions of the adjusted R-squared values are all positively skewed. According to Panel B of Table 1, if sentiment shock is used, the bearish sentiment shock exhibits the higher explanatory power of stock returns, with mean and median values of 8.91% and 7.56%. The bullish sentiment shock ranks the second with mean and median values of 5.63% and 4.67%, respectively. The explanatory power of the neutral shock is still the lowest. The mean and median values of the adjusted R-Squared value are 1.03% and 0.53%, respectively. We can see that, if sentiment shock is used to explain stock returns, the mean and median of the explanatory power are significantly higher than those of the sentiments. Figure 1 exhibits the ranking of the adjusted R-Squared values of the six sentiment measures.

[Insert Figure 1 Here]

We can see that bearish shock exhibits the highest explanatory power to stock returns, followed by bullish shock, bullish sentiment, and bearish sentiment. The explanatory power of neutral sentiment to neutral shock is minimal. The ranking exhibits a clear pattern: sentiment shock exhibits a higher explanatory power than their sentiment counterparts. This pattern holds for all the three sentiment indexes, including the neutral sentiment index. Interestingly, bullish sentiment has a higher explanatory power on stock returns than bearish sentiment. However, the explanatory power of bearish shock is higher than bullish shock. This contrast shows that investors are mainly driven by bullish sentiment, and they are less likely to be affected by a shock in bullish sentiment. On the other hand, although bearish sentiment is not as essential to stocks returns as bullish sentiment, investors are very sensitive to bearish sentiment shock. In the financial market, investor sentiment is more predictable than sentiment shock, and thus the higher explanatory power of sentiment shock is not a complete surprise.

b. Explanatory power across varied markets

To find out whether the pattern holds across different market states, we divide the whole sample period into four subperiods and repeat the tests. From January 2009 to December 2020, each subperiod covers a 3-year period. Table 2 reports the results.

[Insert Table 2 Here]

The first subperiod is from January 2009 to December 2011. As Panel A in Table 2 shows, the mean and median of the adjusted R-squared of the bullish sentiment are 4.73% and 3.03%, much higher than those of the other two sentiment indexes. It is worth noting that bearish sentiment exhibits the lowest explanatory power on stock returns, even lower than that of the neutral sentiment. We do not think this phenomenon surprising, because the stock market of the first subperiod was characterized by a market boom (except for the dip in the first quarter of 2009), so the low explanatory power of bearish sentiment on stock return is expected. In the same subperiod, as Panel B of Table 2 shows, the bullish and bearish shocks exhibit a significantly higher explanatory power on stock returns than the bullish and bearish sentiments. The mean and median adjusted R-squared values of the bullish sentiment shock are 6.81% and 5.48%, respectively. The

two values are 4.37% and 2.88% for the bearish sentiment. During this subperiod, the neutral sentiment shocks have no explanatory power on stock returns at all.

The second subperiod is from January 2012 to December 2014, which is another booming market. In this period, the bullish sentiment has the highest explanatory power on stock returns. The explanatory power of bearish and neutral sentiment is very small. In contrast, as Panel B of Table 2 shows, the mean and median of the adjusted R-squared are 7.44% and 3.78% for the bullish sentiment shock, and they are 5.66% and 2.44% for the bearish sentiment shock. The adjusted R-Squared of neutral sentiment shocks remains very little, with a mean of 0.90% and a median of -0.79%. Once again, we can see that sentiment shock has a higher explanatory power on stock returns than sentiment level.

The third subperiod spans from January 2015 to December 2017. This period is also featured by a market boom. However, more and more investors started to wonder whether stocks were over-valued during this period. Consequently, the market sentiment was not as bullish as it was in the first two subperiods. As Panel A in Table 2 shows, the explanatory power of the bullish sentiment drops significantly during this period. The mean and median of the adjusted R-Squared value are 2.42% and 0.58%. As for the bearish sentiment, the mean and median of the adjusted R-Squared are 4.83% and 1.94%, which are much higher than those of the bullish sentiment. This change suggests that investor sentiment was switching from bullish to bearish. The neutral sentiment does not exhibit any explanatory on stock returns. The switch in the ranking between bullish and bearish sentiments might be caused by the increased sentiment of uncertainty of investors in a prolonged bullish market. According to Panel B in Table 2, for the bullish sentiment shock, the mean and median of the adjusted R-Squared are 2.25% and -0.01%, about the same as those of the bullish sentiment. As for the bearish sentiment shock, the mean and median are 7.92% and 4.68%, which are much higher than those of the bearish sentiment. They are also much higher than the mean and median of the bullish shock. This implies that bearish sentiment dominated bullish sentiment in this period.

The fourth subperiod is from January 2018 to December 2020. During this period, the market was highly volatile with dramatic ups and downs, including the two large drops in late 2018 and in early March of 2020. Given the high volatility of the market, we expect that investor sentiment should play a more important role in explaining stock returns. However, according to

Panel A of Table 2, the mean and median of the adjusted R-Squared of the bullish sentiment are only 0.33% and -0.97%, respectively. The mean and median of the bearish sentiment are 1.11% and -0.51%, which are also rather small. Once again, the neutral sentiment exhibits zero explanatory on stock returns. This result indicates that investor sentiment provides little explanatory power on stock returns in a volatile market. This might be caused by the more frequent swings between bullish and bearish sentiments of investors. Contrary to this finding, Panel B of Table 2 exhibits a quite different result. The mean and median of the adjusted R-Squared of the bullish sentiment shock are 7.73% and 6.21%. As for bearish shock, these two values are as high as 22.19% and 21.37%, respectively. In addition, the explanatory power of neutral sentiment shock is also very high. The mean and median of the adjusted R-Squared of the neutral sentiment shock are 11.17% and 10.23%, which are higher than those of bullish sentiment shock. We also noticed that the means and median values of the adjusted R-Squared are close to each other, and it indicates that the skewness of the distribution of the adjusted R-Squared is much lower than that in the other subperiods.

Table 2 is very informative. Although the stock market in the sample period represents the longest market boom, dominant investor sentiment varies. In the early stage of the booming market, bullish sentiment (shock) demonstrates a high explanatory power on stock returns. The explanatory power of bearish sentiment (shock) is secondary, and the role of neutral sentiment(shock) on stock returns is minimal. However, as the market boom continues, the high explanatory power of bullish sentiment (shock) gives way to the bearish sentiment (shock). When the stock market becomes very uncertain in the last subperiod, the explanatory of all the three sentiments drops to the lowest level. However, during this period, the bearish sentiment shock exhibits an extremely high explanatory power on stock returns, followed by neutral sentiment shock and bullish sentiment shocks. We can infer that investor are more risk-averse in a volatile market, and their sentiment shocks exhibit a much high explanatory power to stock returns. Compared with good news, bad news has a significantly larger impact on investor sentiment and on stock returns.

c. Sentiment beta and sentiment shock beta

Glushkov (2006) develops a stock-by-stock measure of investor sentiment called sentiment beta. It is the coefficient in the time-series regression of individual stock returns on changes in

sentiment. In this section, we examine the sentiment beta and sentiment shock beta, which measure the sensitivity of stock returns to investor sentiment and sentiment shock, respectively. We use the same model in Equation (1), where sentiment beta is the coefficient on investor sentiment, and sentiment shock beta is the coefficient on sentiment shock. Table 3 reports the results.[3]

[Insert Table 3 Here]

We can see from Panel A of Table 3 that, based on the entire sample period, the mean and median of bullish sentiment beta are 0.3264 and 0.2966. Out of the 3,340 sentiment betas, 53.98% are statistically significant at a 5% level. Panel B shows that the mean and median of bullish shock beta are 0.4767 and 0.4445, much higher than those of the bullish sentiment. Moreover, 69.10% of the 3,340 sentiment betas are statistically significant. Table 3 also shows that the bearish sentiment beta has a mean of -0.1796 and a median of -0.2148, and 22.99% of the betas are statistically significant. On the other hand, the mean and median of the bearish shock beta are -0.5123 and -0.4753, and 73.74% of the betas are statistically significant. We can see that stock returns are more sensitive to sentiment stock than to sentiment level, and this holds for both bullish sentiment and bearish sentiment. Table 3 also shows that only 7.76% of the neutral sentiment betas are statistically significant whereas it is 24.37% for neutral sentiment shock.

d. Sentiment (shock) beta across varied markets

We already found that the explanatory power of sentiment and sentiment shock varies across market states. In this section, we test whether market states play a role in sentiment beta and sentiment shock beta. Once again, we divide the entire sample period into four 3-year subperiods. The model setup for sentiment beta is as follows:

$$R_{i,t} = \alpha_i + \beta_i * SL_t + \varepsilon_i \quad (2)$$

[3] The level of statistical significance in the tests is 5%.

Where $R_{i,t}$ is the return of stock i in month t. SL_t stands for the sentiment level of the three AAII sentiment indexes, including the bullish, bearish, and neutral indexes. β_i is the sentiment beta of stock i. Table 4 reports the results.

[Insert Table 4 Here]

From Table 4 we can see that the values of bullish sentiment beta are all positive whereas those of the bearish sentiment are all negative. Although these numbers change across markets, this pattern is not surprising. The interesting part lies in the percentage of statistically significant sentiment beta across the four subperiods. We know that the sample period spans from January 2009 to December 2020. This period is mainly a booming market with some ups and downs, especially in the beginning and toward the end of the sample period. In the first subperiod, the market just recovered from the 2008 financial crisis, and so it was highly bullish. In this period, the percentage of statistically significant bullish sentiment beta is 27.84% while this number is only 8.17% for the bearish sentiment beta. In the second subperiod, this percentage remains the same for bullish sentiment beta, but it rises to 15.39% for the bearish sentiment beta. As the market boom continues, investor sentiment started to change. In the third subperiod, the percentage of statistically significant bullish sentiment beta is 16.41%, whereas this percentage for the bearish sentiment beta rises to 27.93%. This pattern indicates that investors were concerned about what might happen next in the market.

In the last subperiod, the market started having higher volatility, especially in the last quarter of 2018 and in the first quarter of 2020. In this volatile period, the percentage of the statistically significant sentiment beta is quite low for both the bullish and the bearish sentiments. This result indicates that the stock market loses its direction when investors are concerned. In addition, the percentage of statistically significant sentiment beta is 4.94% for bullish sentiment; And for bearish sentiment, it is 8.83%. This suggests that bearish sentiment is more relevant to stock returns in this uncertain period. The neutral sentiment beta does not exhibit any evident pattern, suggesting that the neutral sentiment is not as important as the other two sentiments.

Next, we examine whether sentiment shock beta exhibit any pattern as market state changes. Also, we want to compare the difference in sensitivity to stock returns between sentiment shock beta and sentiment beta. The model setup for sentiment shock beta is as follows:

$$R_{i,t} = \alpha_i + \beta_i * SS_t + \varepsilon_i \quad (3)$$

Where $R_{i,t}$ is the return of stock i in month t. SS_t stands for a sentiment shock of the three AAII sentiment indexes, including the bullish, bearish, and neutral indexes. β_i is the sentiment shock beta of stock i. Table 5 reports the results.

[Insert Table 5 Here]

The first thing we notice from Table 5 is the much higher percentage of statistically significant sentiment shock beta in all the fours subperiods. For the bullish sentiment shock beta, the percentage in the four subperiods are 38.41%, 35.63%, 15.69%, and 42.55%. They are all significantly higher than those of the sentiment beta. This finding also holds for bearish sentiment beta whose percentages of statistically significant beta are 27.04%, 29.34%, 38.74%, and 74.67%, respectively. Moreover, for both the bullish and bearish sentiment shock betas, their mean and median values are also evidently higher than their sentiment beta counterparts. The most striking finding is in the last subperiod, which is highly volatile. In this period, nearly half of the bullish sentiment shock beta is statistically significant, and three quarters of the bearish sentiment shock betas are statistically significant. And 56.59% of the neutral sentiment shock betas are statistically significant. In addition, the values of the sentiment shock beta are well above those of their corresponding sentiment betas. We also find that neutral sentiment shock is way more sensitive to stocks returns than neutral sentiment.

Comparing the results between Table 4 and Table 5, we can see that sentiment shock is more relevant to stock returns than sentiment level. Besides, the importance of sentiment and sentiment shock varies across market states. As for the investor sentiment, bullish sentiment matters more in a bullish market, whereas bearish sentiment is more relevant in a wary market. In a volatile market, the sensitivity of all the three sentiments to stock returns is low. However, in the

same period, sentiment shock becomes highly sensitive to stock returns. Out of the three types of sentiment shocks, bearish sentiment shock shows the highest sensitivity to stock returns. Once again, we can infer that sentiment shock plays a more important role than sentiment in stock returns. And out of the three sentiment shocks, bearish sentiment shock has the highest relevance to stock returns. This finding is consistent with the result in section b.

e. Sentiment, sentiment shock, and market factors

In this section, we examine the relationship between investor sentiment, sentiment shock, and market factors. Specifically, we want to see the explanatory power of market factors on investor sentiment and sentiment shock. If market factors have a higher explanatory power on a sentiment measure, it means that this measure contains more rational elements in regard to stock market. To find out the answer, we use the four Fama-French market factors as the independent variables to test their explanatory power on a sentiment measure. The four market factors are the monthly market excess return, the size factor, the style factor, and the momentum factor. We use the following model in the tests:

$$SI_t = \alpha + \sum_{i=1}^{N} \beta_i \, r_{i,t} + \varepsilon \qquad (4)$$

Where SI_t is the monthly sentiment or sentiment shock; βi is the loading on market factor i, representing the monthly market excess return, the size factor, the style factor, and the momentum factor. Table 6 reports the results.

[Insert Table 6 Here]

Panel A in Table 6 shows that the adjusted R-Squared of the market factors on the bullish and bearish sentiments are almost the same. The former is 8.06% and the latter is 8.03%, followed by the neutral sentiment with an adjusted R-Squared value of 4.57%. We can see that the market return factor is statistically significant if bullish sentiment or bearish sentiment is used as the dependent variable in the regression. The size factor and the style factor are statistically significant if bearish sentiment is the dependent variable. The neutral factor does not exhibit much relationship with the market factors, except for the style factor. We also notice that the alpha is both statistically

and economically significant in all the threes regressions, suggesting that a large portion of investor sentiment cannot be explained by the four market factors. Put it another way, investor sentiment does contain very much emotional elements.

From Panel A we can see that investor sentiments themselves are related to what's going on in the market. However, this relationship is rather weak. Panel B of Table 6 shows that the explanatory power of the four market factors on sentiment shocks is much higher. If bullish shock is used as the dependent variable, the adjusted R-Squared of the regression is 18.52%. This value is 29.41% for the bearish shock. The market return factor is statistically significant if the bullish sentiment shock is the dependent variable. As for the bearish sentiment shock, all but the style factor is statistically significant. The adjusted R-Squared of the neutral sentiment shock is 0.81%, suggesting that the relationship between neutral sentiment shock and market factors is minimal. Figure 2 presents the ranking of the explanatory power of the four Fama-French market factors on sentiments and sentiment shocks. From Table 6 we can see that market factors have a significantly higher explanatory power on sentiment shocks than on sentiment level. This difference indicates that, compared with sentiment, sentiment shock has a closer relationship with market factors. Figure 2 presents the ranking of market factors' explanatory power on investor sentiments and sentiment shocks.

[Insert Figure 2 Here]

Figure 2 shows that the market factors have the highest explanatory power on bearish sentiment shock, followed by bullish sentiment shock, bullish sentiment, bearish sentiment. Again, the relationship between market factors and neutral sentiment or neutral shock is very small. Comparing Figure 1 with Figure 2, we can see a similar pattern in this ranking. We can infer that sentiment shock is much more relevant to stocks returns than investor sentiment. Sentiment itself has weak relationship to the performance of stock market.

Next, we examine the explanatory power of individual market factors on sentiments and sentiment shocks. The purpose of this test is to locate the market factor that sentiment and sentiment shock are most sensitive to. We use the following model in the tests:

$$SI_t = \alpha + \beta_i * F_{i,t} + \varepsilon_i \quad (5)$$

Where SI_t is the monthly sentiment or sentiment shock of the three AAII sentiment indexes, including the bullish, bearish, and neutral indexes. βi is the loading on market factor F_i, representing the monthly market excess return, the size factor, the style factor, and the momentum factor in Fama-French 4-factor model. Table 7 reports the results.

[Insert Table 7 Here]

According to Panel A of Table 7, if bullish sentiment is used as the dependent variable in the regression, the market return factor RMRF has the highest explanatory power, followed by momentum factor MOM, size factor SMB, and style factor HML. The adjusted R-Squared values of the four market factors are 8.97%, 7.85%, 3.60%, and 1.59%, respectively. For the bearish sentiment, the market return factor has the highest adjusted R-Squared of 3.81%, followed by the style factor's 3.08%, and the size factor's 1.17%. The momentum factor has no explanatory power on the bearish sentiment. For the neutral sentiment, we can see that the momentum factor's adjusted R-Squared value is 3.43%, and the other three market factors exhibit no explanatory power to neutral sentiment.

Panel B of Table 7 indicates that the individual market factors have a significantly higher explanatory power on sentiment shocks than on sentiment itself. For example, the adjusted R-Squared of the market return factor is 16.43% if bullish sentiment shock is the dependent variable, followed by momentum factor's 7.68%, size factor's 7.49%, and style factor's 5.45%. As for the bearish sentiment shock, the market return factor has an adjusted R-Squared value of 24.40%, followed by size factor's 12.64% and momentum factor's 12.56%, and style factor's 7.62%. The adjusted R-Squared value of every market factor remains low for neutral sentiment shocks.

This section demonstrates that, compared with investor sentiment, sentiment shock is better explained by Fama-French market factors. This pattern holds for all the three types of investor sentiments, suggesting that sentiment shock has a much closer relationship with stock market. We can also see that the market return factor is the most essential driver for both bullish and bearish

sentiments and their sentiment shocks. It is surprising to see that the momentum factor has a high explanatory power on bullish sentiment, bullish sentiment shock, and bearish sentiment shock. This finding seems to be consistent with Grinblatt, Titman, and Wermers (1995), who found that about 77% of mutual funds were "momentum investors." In addition, we find that the neutral sentiment and sentiment shock hardly exhibit any relationship with the market factors. Overall, we can infer that sentiment shock are more relevant than sentiment level to stock market.

Conclusion

We examined the relationship between stock returns and investor sentiment. This study uses the AAII sentiment indexes, including the bullish sentiment, the bearish sentiment, and the neutral sentiment. We compare investor sentiment with sentiment shock to determine which measure better captures the relationship between sentiment and stock returns.

We find that sentiment shock has a higher explanatory power on stock returns than sentiment level. In addition, the explanatory power of sentiment and sentiment shock on stock returns vary in different market states. In addition, we compared the difference between sentiment beta and sentiment shock beta. We find that sentiment shock beta exhibits a much higher statistical significance than sentiment beta, suggesting that stock returns are more sensitive to sentiment shocks. We also find that, compared with sentiment, sentiment shock has a stronger linkage to the four Fama-French market factors, indicated by the higher explanatory power of the market factors on sentiment shocks. In other words, sentiment shock is more responsive to stock market than sentiment level. Furthermore, we find that the market return factor is the most essential driver for both bullish and bearish sentiments and the corresponding sentiment shocks. Momentum factor and size factor also play a significant role in explaining sentiment and sentiment shock.

We conclude that sentiment shock is a better sentiment measure regarding the relationship between investor sentiment and stock returns. Stock returns are more sensitive to sentiment shock than sentiment level. Since sentiment shock is more unexpected than sentiment level, it captures more information about investors' decision-making process related to stock returns.

References

Amihud, Yakov and Ruslan Goyenko, 2013. "Mutual Fund's R-super-2 as Predictor of Performance," *Review of Financial Studies*, Vol. 26, No. 3, pp. 667-694.

Baker, Malcolm, Jeffrey Wurgler, 2006. “Investor Sentiment and the Cross-Section of Stock Returns”, Journal of Finance, Vol. 61, Issue 4, pp. 1645–1680.

Beaumont, Rob , Marco van Daele, Bart Frijns, Thorsten Lehnert, Aline Muller, 2008. "Investor sentiment, mutual fund flows and its impact on returns and volatility", *Managerial Finance*, Vol. 34, No. 11, pp. 772-785.

Brown, Gregory W. & Cliff, Michael T., 2004. "Investor sentiment and the near-term stock market," *Journal of Empirical Finance*, Elsevier, Vol. 11, No. 1, pp. 1-27.

Qiang Bu, “Are all the sentiment measures the same?” *Journal of Behavioral Finance,* Forthcoming. Published online: July 24,2021. https://www.tandfonline.com/doi/abs/10.1080/15427560.2021.1949718

Bu, Qiang, and Jeffrey Forrest. 2020. “Comparing sentiment measures in mutual fund performance,” *International Journal of Managerial Finance*, Forthcoming: https://www.emerald.com/insight/content/doi/10.1108/IJMF-02-2020-0092/full/html

Chu, Liya, Qianqian Du, and Jun Tu, 2017. Purging Investor Sentiment Index from Too Much Fundamental Information, *Working Paper.*

Cipollini, A.P.L., and A. Manzini, 2007, Can the VIX signal market's direction? An asymmetric dynamic strategy, working paper.

Clarke, Roger G. , and Meir Statman, 1998. Bullish or Bearish? *Financial Analysts Journal,* Volume 54, No. 3, pp. 63-72.

De Bondt, Werner, F.M., and Thaler, Richard, 1985. "Does the Stock Market Overreact?," *Journal of Finance*, Vol. 40, No. 3, pp. 793-805.

De Bondt, Werner, F.M., 1993. “Betting on trends: Intuitive forecasts of financial risk and return,” *International Journal of Forecasting*, Vol. 9, Issue 3, pp. 355-371.

Dong, Feng, and John Doukas, 2018. Fund Management Skill and Noise Trading, *Journal of Portfolio Management*, Vol. 44, Issue 5, pp. 113-124.

Feldman, Todd. 2010, A More Predictive Index of Market Sentiment, *Journal of Behavioral Finance,* Vol. 11, No. 4, 211-223.

Fleming, Jeff, Barbara Ostdiek, and Robert Whaley, 1995, Predicting stock market volatility: a new measure, *Journal of Futures Markets,* Vol. 15, No. 3, pp. 265-302.

Feldman, Todd, and Shuming Liu, 2017. Contagious Investor Sentiment and International Markets, *Journal of Portfolio Management,* Vol. 43, Vol. 4, pp. 125-136.

Grinblatt, Mark, Sheridan Titman and Russ Wermers, 1995. Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior. *American Economic Review,* Vol. *85*, No. 5, 1088-1105.

Goetzmann, William N., Massimo Massa, K. Geert Rouwenhorst, 2000. Behavioral Factors in Mutual Fund Flows. Yale Working Paper.

Glushkov, Denys, 2006. Sentiment Beta. Working Paper. Available at SSRN 862444, 2006.

Kahneman, Daniel, and Amos Tversky, Prospect Theory: An Analysis of Decision under Risk, Econometrica, Vol. 47, No. 2, pp. 263-291.

Lemmon, Michael, and Evgenia Portniaguina, 2006. "Consumer Confidence and Asset Prices: Some Empirical Evidence," *Review of Financial Studies*, Vol. 19, Issue 4, pp. 1499-1529.

Otoo, Maria Ward, 1999. "Consumer sentiment and the stock market," Finance and Economics Discussion Series 1999-60, Board of Governors of the Federal Reserve System (U.S.).

Qiu, Lily, and Ivo Welch, 2004. Investor Sentiment Measures, NBER Working Paper 1079.

Sibley, Steven E., Yanchu Wang, Yuhang Xing, Xiaoyan Zhang, 2016. "The information content of the sentiment index," *Journal of Banking & Finance*, Volume 62, No. 1, pp. 164-179.

Statman, Meir, 2011. "Investor Sentiment, Stock Characteristics, and Returns," *Journal of Portfolio Management*, Vol. 37, No. 3, pp. 54-61.

Solt, Michael E., and Meir Statman. 1988. "How Useful is the Sentiment Index?" *Financial Analyst Journal*, Vol. 44, No. 5, pp. 45-55.

Wang, Yaw-Huei, Aneel Keswani and Stephen J. Taylor, 2006. "The relationships between sentiment, returns and volatility," *International Journal of Forecasting*, Vol. 22, No. 1, pp. 109-123.

Wulfmeyer, Sina. 2016. Irrational Mutual Fund Managers: Explaining Differences in Their Behavior, *Journal of behavioral finance*, Volume 17, No. 2, 99-123.

Table 1: Explanatory power of investor sentiment on stock returns

Table 1 reports the explanatory power of investor sentiment on stock returns. The explanatory power is proxied by the adjusted R-Squared values of a model with the following setup:

$$R_{i,t} = \alpha_i + \beta_i * S_t + \varepsilon_i$$

Where $R_{i,t}$ is the return of stock i in month t. S_t stands for a sentiment or sentiment shock of the three AAII sentiment indexes, including the bullish, bearish, and neutral sentiment indexes. The stock sample includes 3,340 stocks, and the test results are based on monthly return data. The sample period spans a 12-year period from January 2009 to December 2020.

Panel A: Sentiment			
	Bullish	Bearish	Neutral
Mean	2.96%	1.59%	0.25%
Median	2.31%	0.75%	-0.12%
Panel B: Sentiment shock			
	Bullish Shock	Bearish Shock	Neutral Shock
Mean	5.63%	8.91%	1.03%
Median	4.67%	7.56%	0.53%

Figure 1: Ranking of the explanatory power of the sentiment and the sentiment shock

Figure 1 presents the ranking of the explanatory power of investor sentiment on stock returns. The explanatory power is proxied by the adjusted R-Squared values of a model with the following setup:

$$R_{i,t} = \alpha_i + \beta_i * S_t + \varepsilon_i$$

Where $R_{i,t}$ is the return of stock i in month t. S_t stands for the sentiment or the sentiment shock of the three AAII sentiment indexes, including the bullish, bearish, and neutral sentiment indexes. The stock sample includes 3,340 stocks, and the test results are based on monthly return data. The sample period spans a 12-year period from January 2009 to December 2020.

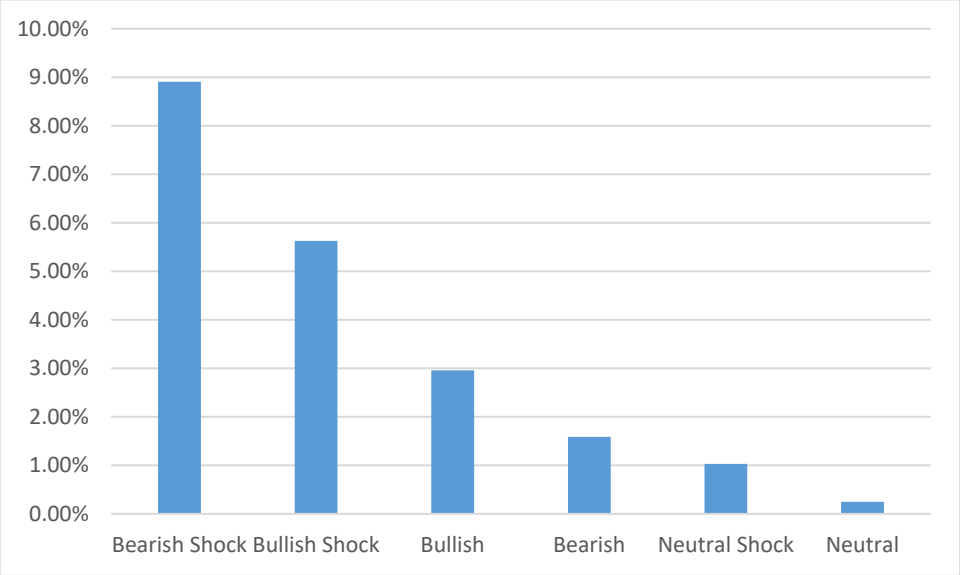


Table 2: Explanatory power of investor sentiment indexes on stock returns

Table 2 reports the mean and median of the adjusted R-squared values of investor sentiment in four subperiods of the whole sample period. The model setup is as follows:

$$R_{i,t} = \alpha_i + \beta_i * S_t + \varepsilon_i$$

Where $R_{i,t}$ is the return of stock i in month t. S_t stands for the sentiment or the sentiment shock of the three AAII sentiment indexes, including the bullish, bearish, and neutral indexes. The stock sample includes 3,340 stocks, and the test results are based on monthly return data. The whole sample period spans a 12-year period from January 2009 to December 2020.

Panel A: Sentiment				
	Stats	Bullish	Bearish	Neutral
01/2009-12/2011	Mean	4.73%	0.92%	1.97%
	Median	3.03%	-0.88%	0.512%
01/2012-12/2014	Mean	5.25%	2.15%	0.66%
	Median	1.97%	-0.17%	-0.86%
01/2015-12/2017	Mean	2.42%	4.83%	-1.14%
	Median	0.58%	1.94%	-2.15%
01/2018-12/2020	Mean	0.33%	1.11%	-0.18%
	Median	-0.97%	-0.51%	-1.44%
Panel B: Sentiment shock				
	Stats	Bullish Shock	Bearish Shock	Neutral Shock
01/2009-12/2011	Mean	6.81%	4.37%	-0.31%
	Median	5.48%	2.88%	-1.80%
01/2012-12/2014	Mean	7.44%	5.66%	0.90%
	Median	3.78%	2.44%	-0.79%
01/2015-12/2017	Mean	2.25%	7.92%	2.02%
	Median	-0.01%	4.68%	-0.34%
01/2018-12/2020	Mean	7.73%	22.19%	11.17%
	Median	6.21%	21.37%	10.23%

Table 3: Sentiment beta and sentiment shock beta

Table 3 reports the percentage of the sentiment beta statistically significant at 5%. It also presents the mean and median adjusted R-squared values of investor sentiment on stock returns. The model setup is as follows:

$$R_{i,t} = \alpha_i + \beta * S_t + \varepsilon_i$$

Where $R_{i,t}$ is the return of stock i in month t. S_t stands for a sentiment or sentiment shock of the three AAII sentiment indexes, including the bullish, bearish, and neutral indexes. β is the sentiment beta on either sentiment or sentiment shock. The stock sample includes 3,340 stocks, and the test results are based on monthly return data. The whole sample period spans from January 2009 to December 2020.

Panel A: Sentiment			
	Bullish	Bearish	Neutral
Statistically Significant (%)	53.98%	22.994%	7.755%
Mean	0.3264	-0.1796	-0.4053
Median	0.2966	-0.2148	-0.3822
Panel B: Sentiment shock			
	Bullish Shock	Bearish Shock	Neutral Shock
Statistically Significant (%)	69.102%	73.743%	24.37%
Mean	0.4767	-0.5123	0.4764
Median	0.4445	-0.4753	0.4241

Table 4: Sentiment beta across market states

Table 4 reports the percentage of the sentiment beta statistically significant at 5% in the four subperiods of the whole sample period. It also presents the mean and median adjusted R-squared values of investor sentiment on stock returns. The model setup is as follows:

$$R_{i,t} = \alpha_i + \beta_i * SL_t + \varepsilon_i$$

Where $R_{i,t}$ is the return of stock i in month t. SL_t stands for the sentiment level of the three AAII sentiment indexes, including the bullish, bearish, and neutral indexes. β_i is the sentiment beta of a stock. The stock sample includes 3,340 stocks, and the test results are based on monthly return data. The whole sample period spans from January 2009 to December 2020.

Period	Statistical significance	Bullish	Bearish	Neutral
01/2009-12/2011	Statistical Significant (%)	27.84%	8.17%	11.14%
	Mean	0.5582	-0.2186	-1.3354
	Median	0.5532	-0.3083	-1.0840
01/2012-12/2014	Statistical Significant (%)	27.67%	15.39%	5.81%
	Mean	0.4414	-0.4229	-0.6869
	Median	0.3870	-0.3770	-0.6108
01/2015-12/2017	Statistical Significant (%)	16.41%	27.93%	1.49%
	Mean	0.4314	-0.5894	-0.4292
	Median	0.3272	-0.5287	-0.6150
01/2018-12/2020	Statistical Significant (%)	4.94%	8.83%	3.35%
	Mean	0.7377	-0.4018	0.3333
	Median	0.7433	-0.4549	0.4982

Table 5: Sentiment shock beta across market states

Table 5 reports the percentage of the sentiment beta statistically significant at 5% in the four subperiods of the whole sample period. It also presents the mean and median adjusted R-squared values of investor sentiment on stock returns. The model setup is as follows:

$$R_{i,t} = \alpha_i + \beta_i * SS_t + \varepsilon_i$$

Where $R_{i,t}$ is the return of stock i in month t. SS_t stands for a sentiment shock of the three AAII sentiment indexes, including the bullish, bearish, and neutral indexes. β_i is the sentiment shock beta of stock i. The stock sample includes 3,340 stocks, and the test results are based on monthly return data. The whole sample period spans from January 2009 to December 2020.

Period	Statistical significance	Bullish Shock	Bearish Shock	Neutral Shock
01/2009-12/2011	Statistical Significant (%)	38.41%	27.04%	4.52%
	Mean	0.8467	-0.8019	-1.2427
	Median	0.7543	-0.6847	-1.2904
01/2012-12/2014	Statistical Significant (%)	35.63%	29.34%	6.77%
	Mean	0.4195	-0.4612	-0.8538
	Median	0.3758	-0.4137	-0.7512
01/2015-12/2017	Statistical Significant (%)	15.69%	38.74%	14.82%
	Mean	0.6586	-0.6253	0.5095
	Median	0.5526	-0.5451	0.4489
01/2018-12/2020	Statistical Significant (%)	42.55%	74.67%	56.59%
	Mean	0.9016	-0.9391	1.0393
	Median	0.8046	-0.8701	0.9487

Table 6: The explanatory power of market factors on investor sentiment

Table 6 reports the average adjusted R-squared of the regressions, which use individual market factors to explain a sentiment or sentiment shock. We use the following model in the tests:

$$S_t = \alpha + \sum_{i=1}^{N} \beta_i \, r_{i,t} + \varepsilon$$

Where S_t is the monthly sentiment or sentiment shock; βi is the loading on market factor i, representing the monthly market excess return, the size factor, the style factor, and the momentum factor. The stock sample includes 3,340 stocks, and the test results are based on monthly return data. The sample period spans a 12-year period from January 2009 to December 2020.

Sentiment index	Incept	RMRF	SMB	HML	MOM	Adj. R-sq
Panel A: Sentiment						
Bullish	0.355 (63.10)***	0.004 (2.72)***	0.003 (1.13)	0.001 (0.37)	-0.000 (-0.11)	8.06%
Bearish	0.324 (54.47)***	-0.003 (-2.10)**	-0.002 (-0.83)	-0.005 (-2.38)**	-0.004 (-2.58)**	8.03%
Neutral	0.320 (51.79)***	-0.001 (-0.50)	-0.001 (-0.25)	0.004 (2.15)**	0.004 (2.85)	4.57%
Panel B: Sentiment shock						
Bullish Shock	-0.004 (-0.89)	0.004 (3.33)***	0.003 (1.60)	0.001 (0.70)	-0.001 (-1.32)	18.52%
Bearish Shock	0.005 (1.18)	-0.005 (-4.23)***	-0.004 (-2.37)**	-0.001 (-0.62)	0.002 (1.99)**	29.41%
Neutral Shock	-0.001 (-0.34)	0.001 (1.01)	0.001 (0.91)	-0.000 (-0.13)	-0.001 (-0.81)	0.81%

Figure 2: Ranking of the Explanatory power of market factors on sentiment

Table 1 reports the ranking of the explanatory power of investor sentiment on stock returns. The explanatory power is proxied by the adjusted R-Squared values of a model. We use the following model in the tests:

$$SI_t = \alpha + \sum_{i=1}^{N} \beta_i \, r_{i,t} + \varepsilon$$

Where SI_t is the monthly sentiment or sentiment shock of the three AAII sentiment indexes, including the bullish, bearish, and neutral indexes. βi is the loading on market factor i, representing the monthly market excess return, the size factor, the style factor, and the momentum factor in Fama-French 4-factor model. The stock sample includes 3,340 stocks, and the test results are based on monthly return data. The sample period spans a 12-year period from January 2009 to December 2020.

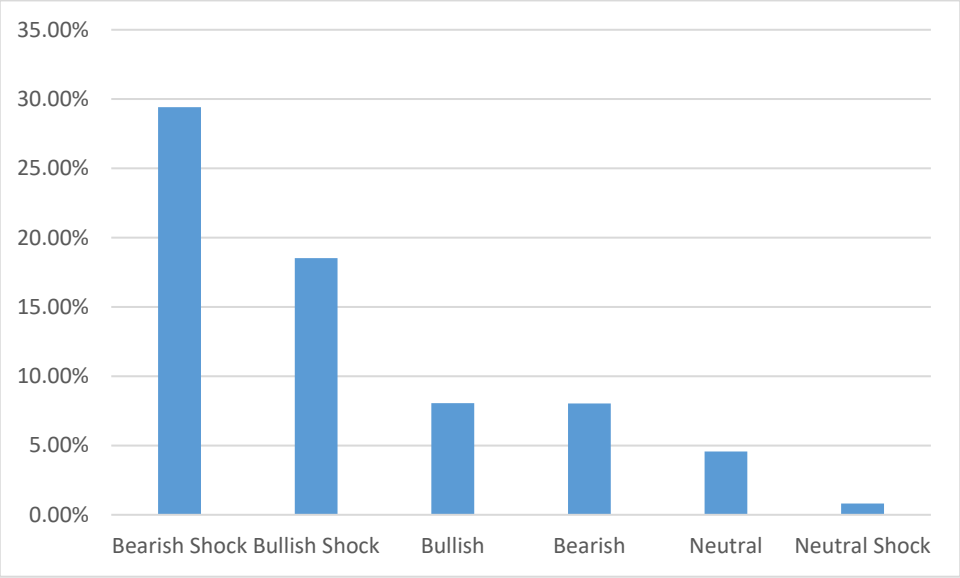


Table 7: Explanatory power of individual market factors on investor sentiment

Table 7 reports the average adjusted R-squared of the regressions, which use the individual market factors in Fama-French 4-factor model to explain investor sentiment. We use the following model in the tests:

$$S_t = \alpha + \beta_i * F_{i,t} + \varepsilon_i$$

Where S_t is the monthly sentiment or sentiment shock of the three AAII sentiment indexes, including the bullish, bearish, and neutral indexes. βi is the loading on a market factor F_i, representing the monthly market excess return, the size factor, the style factor, and the momentum factor in Fama-French 4-factor model. The stock sample includes 3,340 stocks, and the test results are based on monthly return data. The sample period spans a 12-year period from January 2009 to December 2020.

Panel A: Sentiment level			
	Bullish	**Bearish**	**Neutral**
RMRF	8.97%	3.81%	0.08%
SMB	3.60%	1.17%	-0.29%
HML	1.59%	3.08%	-0.39%
MOM	7.85%	-0.47%	3.43%
Panel B: Sentiment shock level			
	Bullish Shock	**Bearish Shock**	**Neutral Shock**
RMRF	16.43%	24.40%	1.75%
SMB	7.49%	12.64%	1.17%
HML	5.45%	7.62%	-0.12%
MOM	7.68%	12.56%	0.99%

Abstract

Using extreme Bitcoin returns as attention shocks, this study examines whether investors exhibit rational inattention to the stock market. During attention shocks, investors allocate more attention to market-level information and less attention to firm-specific information due to limited attention. This "category-learning" behavior results in an increased stock return co-movement with the market. Our empirical examination provides supporting evidence that extreme Bitcoin returns are associated with larger stock returns co-movement. Further analysis shows that the level of rational inattention attributed to the extreme Bitcoin returns is more profound during the COVID-19 pandemic.

Keywords: Extreme Bitcoin Returns, Attention Allocation, Stock Return Co-movement, COVID-19

JEL classification: G12, G14

Highlights:

- Extreme Bitcoin returns are associated with an increase in the stock return co-movement with the market.
- Findings support the "category-learning" behavior that investors pay rational inattention to firm-specific information by focusing more on market-level information on extreme Bitcoin return days.
- COVID-19 amplifies the rational inattention attributed to extreme Bitcoin returns.

Extreme Bitcoin Returns and Investors Inattention to the Stock Market:

Evidence from Stock Return Co-movement, Before and During COVID-19

1. Introduction

Being the largest capitalization in the cryptocurrency market, Bitcoin has received great interest from the media, investors, and scholars (Urquhart, 2018). Bitcoin is a virtual currency that carries the characteristics of a speculative asset, and shares market behaviors that are different from many conventional assets (Baur, Hong, & Lee, 2018; Klein, Thu, & Walther, 2018). Compared with the traditional stock market, Bitcoin is far more volatile and has a greater magnitude of extreme returns (Núñez, Contreras-Valdez, & Franco-Ruiz, 2019). While a growing body of research shows that investors' attention explains Bitcoin prices behavior (Choi, 2021; Dastgir, Demir, Downing, Gozgor, & Lau, 2019; Philippas, Rjiba, Guesmi, & Goutte, 2019; Zhang, Lu, Tao, & Wang, 2021), few studies place an exclusive focus on the attention allocation across the Bitcoin market and the stock market. This study aims to fill this research gap by examining how extreme Bitcoin returns, as attention shocks, influence the investors' behavior in the stock market. In addition, this study further investigates this issue in the context of the COVID-19 pandemic.

According to the rational inattention framework (Sims, 2003), investors with limited attention can engage in "category-learning", in which they focus more on market-wide information and pay less attention to costly firm-specific information under exogenous attention shocks (Peng & Xiong, 2006). Therefore, this rational inattention is generally associated with a higher level of stock return co-movement with the market during attention shocks (Hu, Li, Goodell, & Shen, 2021; Huang, Huang, & Lin, 2019). In this study, we hypothesize that extreme Bitcoin returns are associated with a larger level of stock return co-movement, as they are an

instantiation of exogenous attention shocks that attract away investors' attention away from the stock market. To test this hypothesis, we focus on all the stocks that are traded in the US market and examine the individual stock's return co-movement on extreme Bitcoin return days versus non-extreme Bitcoin return days. This is the first study that assesses how Bitcoin market influences the investors' behavior in the stock market through the lens of attention allocation at the individual stock level. [1] In addition, inspired by the current studies on the role of COVID-19 on investors' attention (Zhang, Hu, & Ji, 2020) as well as the cryptocurrency market (Guo, Lu, & Wei, 2021; Huang, Duan, & Mishra, 2021; Vidal-Tomás, 2021), we perform a further examination on how rational inattention attributed by the extreme Bitcoin returns varies during the pre-COVID and COVID periods. We expect that the outbreak of COVID-19, as a systematic shock, distracts investors and further moderates the level of rational inattention attributed to extreme Bitcoin returns.

Following Huang et al. (2019), we document cross-sectional evidence that extreme Bitcoin returns are associated with a higher level of stock return co-movement, which suggests the presence of attention distraction based on the "category-learning" assumption. By decomposing the change in stock return co-movement (correlation coefficients), we show that return covariance is the major contributor to the change, followed by market return volatility and stock return volatility. Our empirical results based on the sub-samples in both the pre-COVID-19

[1] Hu, Li, and Shen (2020) also use extreme Bitcoin returns as attention shocks. Unlike ours, they find that extreme Bitcoin returns are associated with a lower level of return co-movement between stock market indices and a composite global market index, which cannot be explained by the "category-learning". As the key assumption in "category learning" is investors have limited attention, which generally does not hold at the global indices level, attributing their findings to attention allocation is not appropriate. In addition, their study has a limited sample size (only 34 indices), which also causes issues when determining the global market index and assessing the co-movement.

period and COVID period reveal that investors' rational inattention on extreme Bitcoin returns days is amplified during the COVID-19 pandemic. The findings advance our understanding of attention allocation across the cryptocurrency market and the stock market.

The remainder of this paper is structured as follows. Section 2 describes data and justifies the proxy of extreme Bitcoin returns for attention shocks. Section 3 presents the results of our main analysis. Section 4 examines how the main findings vary during the pre-COVID period and COVID period. Section 5 concludes the paper.

2. The Data

We obtain historical data return from the CRSP database and historical Bitcoin prices from CoinMarketCap (www.coinmarketcap.com). Our sample includes 5,981 US common stocks and covers the period between Jan-1, 2015 and Dec-31, 2021. We calculate the Bitcoin daily return as $R_t = \frac{Close_t}{Close_{t-1}} - 1$, where $Close_t$ is the Bitcoin daily close price on day *t*.

To justify why extreme Bitcoin returns are a suitable proxy for attention shocks, we collect google search volume of interest (GSVI) from Google Trend, as it is often used as a superior measure for retail investors' attention (Da, Engelberg, & Gao, 2011).[2] We calculate the daily change of search interest in Bitcoin as $\Delta GSVI_t = \ln(GSVI_t) - \ln(GSVI_{t-1})$, where $GSVI_t$ represents the web search interest with the keyword "Bitcoin" from the US on day *t*. Figure 1 presents the relationship between the Bitcoin daily returns and the change of daily search interest

[2] Following Zhang et al. (2021), we obtain the daily GSVI in two steps. In step 1, we download monthly trend covering the full sample period. In step 2, we download daily trend for each month and merge them back to the monthly trend file. We use the monthly trend hits as adjustment factor to retrieve daily trend. Since this study focuses on US stocks, we chose the search location to be "US". When the search of interest is displayed as "<1", we replace it with 0.5.

in “Bitcoin”. This figure shows that both extremely positive and negative Bitcoin returns are associated with a large increase in the search interest, lending support to the appropriateness of using extreme Bitcoin returns as attention shocks.

<Figure 1 here>

To further illustrate the extreme Bitcoin returns, we compare the returns of Bitcoin with those of the S&P500 index. We calculate the top/bottom p^{th} percentiles of daily returns in the previous 100-day rolling window to define “extreme returns” for both Bitcoin and S&P500 index. This rolling estimation helps us avoid the “look-ahead” bias in defining the extreme returns and attend to the most recent returns which reflects the time-varying magnitude of attention shocks. The comparison is reported in Table 1. The average extreme returns of Bitcoin, regardless of different rolling percentiles, are considerably larger than those in the stock market across the years. It is noteworthy that even the magnitude of Bitcoin returns defined by the top/bottom 20% rolling percentiles still exceeds that of the S&P500 index defined by the top/bottom 1% rolling percentiles.

<Table 1 here>

Zhang et al. (2021) find that the causal relationship between Bitcoin returns and investors’ attention is time-varying, which suggests that attention distraction can occur before, during, or after the extreme returns. To identify the most suitable event window around the extreme Bitcoin returns to proxy for attention shock events, we calculate the abnormal search volume $AbGSVI_{(start,end)}$ around extreme Bitcoin returns.

$$AbGSVI_{(start,end)} = \frac{\frac{1}{end - start + 1}\sum_{t=start}^{t=end} GSVI_t}{\frac{1}{20}\sum_{t=start-20}^{t=start-1} GSVI_t} - 1 \quad (1)$$

where *start* and *end* denote the starting and ending dates for the period of interest;

$\frac{1}{end-start+1}\sum_{t=start}^{t=end} GSVI_t$ is the average search volume during that period;

$\frac{1}{20}\sum_{t=start-20}^{t=start-1} GSVI_t$ is the average search volume in the previous 20-day window, which is used to proxy for normal search volume.[3] We evaluate $AbGSVI$ at three periods: i) the event period, in which consecutive extreme returns (or one single extreme Bitcoin return) occur(s); ii) the pre-event period, [-3, -1] three-day window before the first extreme Bitcoin return day; iii) post-event period, [1, 3] three-day window after the last extreme Bitcoin return day.

<Table 2 here>

Table 2 reports the average of ($AbGSVI$) around extreme Bitcoin return days. As expected, we find significant abnormal search volumes in all three periods of interest. The $AbGSVI$ in the pre-event window appears to be lower than it is in both the during-event window and post-event window. The peak level of attention is consistently observed during the extreme Bitcoin return days. Therefore, we decide to only use the extreme Bitcoin return days as the attention shock days in the following analysis.

3. Empirical Analysis and Results

3.1 Change in the Stock Return Co-movement

Following Huang et al. (2019), we use two measures for stock return co-movement: i) Pearson correlation coefficient between the stock return and the market return (Anton & Polk,

[3] Previous studies using weekly search interest (Da et al., 2011) generally define normal search volume as the median of GSVI in the previous seven weeks. We use the average GSVI in the previous 20-day window to account for larger variations at the daily level. Our results remain qualitative unchanged when the estimation window is adjusted to 30-day, 40-day, and 60-day or the median of the GSVI is used.

2014; Peng & Xiong, 2006), and ii) the adjusted R^2 obtained from the CAPM model (Barberis, Shleifer, & Wurgler, 2005; Durnev, Morck, Yeung, & Zarowin, 2003; Morck, Yeung, & Yu, 2000). For each stock *i*, we calculate the Pearson correlation between the stock return and the market return on extreme Bitcoin return days ($Cor_i^{BtcExtRet}$) and on non-extreme Bitcoin return days ($Cor_i^{NonBtcExtRet}$), respectively. Similarly, the respective adjusted R^2 is obtained by fitting the following regression on extreme Bitcoin return days ($Adj\ R_i^{2\ BtcExtRet}$) and on non-extreme Bitcoin return days ($Adj\ R_i^{2\ NonBtcExtRet}$).

$$Ret_{i,t} = \alpha_i + \beta_i MktRet_i + \epsilon_{i,t} \quad (2)$$

where $Ret_{i,t}$ is the daily excess stock return and $MktRet_i$ is the daily excess market return.[4] To obtain reliable co-movement measures, we require that each stock should have at least 15 daily observations on both extreme Bitcoin return days and non-extreme Bitcoin return days. We use $Cor_i^{BtcExtRet} - Cor_i^{NonBtcExtRet}$ and $Adj\ R_i^{2\ BtcExtRet} - Adj\ R_i^{2\ NonBtcExtRet}$ to measure the change in the stock co-movement, which reflects investors' attention allocated from firm-specific information to market-wide information. In addition, we perform a two-sample correlation coefficient test (Jennrich, 1970) for each stock to examine whether $Cor_i^{BtcExtRet} - Cor_i^{NonBtcExtRet}$ is significantly positive or negative at 95% confidence level.

<Table 3>

We report the results in Table 3. When Bitcoin returns exceed the top/bottom 1% (20%) rolling percentiles, the net change in the correlation coefficient is 0.131 (0.063), and the

[4] The market daily return is obtained from the Fama-French market factors through http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

difference in the adjusted R^2 is 0.214 (0.048).[5] In addition, columns (7)-(9) reveal that more than 50% of the stocks exhibit a significantly positive change in the correlation coefficient on extreme Bitcoin return days and that only fewer than 10% of the stocks receive attention enhancement during Bitcoin shocks, i.e., significantly negative change in the correlation coefficients on extreme Bitcoin return days. These results together suggest that extreme Bitcoin returns distract investors' attention paid to firm-specific information in the stock market, which supports our hypothesis.

3.2 Decomposition of Change in the Correlation Coefficients

To further examine the key drivers of the change in stock return co-movement, we follow Huang et al. (2019) to decompose the change in the correlation coefficient into three components: return covariance, stock return volatility, and market return volatility. We then assess the percentage contribution from each component.

$$1 = \overbrace{\left[\frac{\ln \frac{Cov_i^{BtcExtRet}}{Cov_i^{NonBtcExtRet}}}{\ln \frac{Cor_i^{BtcExtRet}}{Cor_i^{NonBtcExtRet}}}\right]}^{A_i} + \overbrace{\left[-\frac{\ln \frac{RetStd_i^{BtcExtRet}}{RetStd_i^{NonBtcExtRet}}}{\ln \frac{Cor_i^{BtcExtRet}}{Cor_i^{NonBtcExtRet}}}\right]}^{B_i} + \overbrace{\left[-\frac{\ln \frac{MktStd_i^{BtcExtRet}}{MktStd_i^{NonBtcExtRet}}}{\ln \frac{Cor_i^{BtcExtRet}}{Cor_i^{NonBtcExtRet}}}\right]}^{C_i} \quad (3)$$

where A_i, B_i, and C_i represent the respective percentage contributions from covariance, stock return volatility, and market return volatility for each stock *i*. Cross-sectional tests are conducted to examine whether the percentage contributions, measured in mean and median, are statistically different from zero.[6]

[5] To provide a comparable understanding on how these values mean, we compare them with findings by Huang et al. (2019). Huang et al. (2019) report the change in the correlation coefficient to be 0.018 and change in the correlation coefficient to be 0.021, when using the large jackpot of Taiwanese nationwide lotteries to proxy for exogenous attention shocks.

[6] Following Huang et al. (2019), each component is winsorized independently at 97.5% and 2.5% levels to mitigate the effect of extreme values.

<Table 4 here>

Table 4 presents the results of correlation coefficient decomposition. Consistent with Huang et al. (2019), we find that the covariance component is the primary contributor to the change in correlation coefficient on extreme Bitcoin return days, which further supports the rational inattention hypothesis. Yet, unlike the findings by Huang et al. (2019), both the stock return volatility and market return volatilities components are significant contributors to the change in the correlation coefficient.[7] These results reveal that due to the volatility spillovers across the Bitcoin market and the stock market, attention distraction by the extreme Bitcoin returns is more complicated than the rational inattention by other types of exogenous attention shocks such as large jackpot lottery (Hu et al., 2021; Huang et al., 2019).

In summary, the stock return co-movement analysis produces strong evidence in support of our hypothesis that extreme Bitcoin returns attract investors' attention away from the stock market. Specifically, we find that such an attention distraction is the main driver of the increased change in stock return co-movement on extreme Bitcoin return days. These findings remain consistent and robust when using different rolling percentiles of Bitcoin returns to define the attention shocks.

4. Further Examination in the Pre-COVID and COVID Periods

Our data sample shares a two-year overlap with another attention distraction event, the COVID-19 pandemic. Recent research shows that the COVID-19 pandemic affects the behavior of both the stock market (Ozkan, 2021) and cryptocurrency markets (Huang et al., 2021; Vidal-Tomás, 2021) as well as investors' attention (Smales, 2020, 2021). Specifically, Bitcoin and the

[7] Huang et al. (2019) report 138% contribution from covariance and -26% contribution from market return volatility, and almost no contribution from stock return volatility.

stock market tend to exhibit larger co-integration and return spillovers during the pandemic (Conlon & McGee, 2020; Nguyen, 2021), which is echoed by our findings documented in Section 3.2. To further examine whether our conclusions remain consistent when accounting for the impact of COVID-19, we repeat the main analysis of the change in the correlation coefficient and $Adj\ R^2$ for each stock (Table 3 in Section 3.1) based on the pre-COVID sub-sample (2015-2019), and the COVID sub-sample (2020-2021), respectively.[8]

<Table 5 here>

We report the results in Table 5. Three observations are worth noting. First, consistent with the results of Table 3, extreme Bitcoin returns are associated with increases in the stock return co-movement in both sub-samples. Second, the comparison of columns 1 and 4 across Panels A and Panel B reveals that the magnitude of the stock return co-movement during COVID-19 is more intensive than that in the pre-COVID period, suggesting that COVID-19 also distracts investors' attention away from the stock market. Lastly, columns 3 and 6 across two panels show that the level of attention distraction due to extreme Bitcoin returns is more profound during the COVID-19 period, which suggests that COVID-19 positively moderates the investors' rational inattention induced by extreme Bitcoin returns.

To conduct a robustness check for the moderating impact of COVID-19, we conduct the following regression analysis.

$$\begin{aligned} Cor_i &= D_i^{COVID} + D_i^{BtcExtRet} + D_i^{BtcExtRet} \times D_i^{COVID} + \epsilon_i \\ Adj\ R_i^2 &= D_i^{COVID} + D_i^{BtcExtRet} + D_i^{BtcExtRet} \times D_i^{COVID} + \epsilon_i \end{aligned} \quad (4)$$

[8] We also use a two-year pre-COVID sub-sample (2018-2019) to match the lengths of the pre-COVID sub-sample and COVID sub-sample. The findings remain qualitatively unchanged.

where D_i^{COVID} and $D_i^{BtcExtRet}$ are two dummies. $D_i^{COVID} = 1$ indicates that the target variables, Cor_i and $Adj\ R_i^2$, are calculated based on the COVID-19 sub-sample; otherwise, they are calculated based on the pre-COVID sub-sample. $D_i^{BtcExtRet}$ =1 indicates that the target variables, Cor_i and $Adj\ R_i^2$, are calculated on extreme Bitcoin return days; otherwise, they are calculated on non-extreme Bitcoin return days. We control for stock fixed-effects and perform standard error clustering by Fama-French 49 industry classification.

<Table 6 here>

The regression results are reported in Table 6. Consistent with Table 5, the significant positive coefficients of D^{COVID} show that COVID-19 is an attention distraction to investors, as it induces a higher level of stock return co-movement. More importantly, the interaction terms, $D^{BtcExtRet} \times D^{COVID}$, are consistently significantly positive across all models, which confirms that the outbreak of COVID-19 amplifies investors' attention distraction in the event of extreme Bitcoin returns.

Finally, we repeat the correlation coefficient decomposition (Table 4 in Section 3.2) based on the pre-COVID and COVID sub-samples. Table 7 presents the results. Consistent with the findings in Table 4, the covariance component is still the primary driver of stock return co-movement in both sub-samples. Across both panels, the contributions from stock return and market return volatilities components are remarkably larger during the COVID period. This observation agrees with the recent findings of the strengthened contagion effect between Bitcoin and the stock market during the COVID-19 pandemic (Guo et al., 2021).

5. Conclusion

This paper makes two key contributions to the literature. First, to our best knowledge, ours is the first study that examines investors' "category-learning" behavior during extreme Bitcoin returns at the stock level. We document strong evidence that extreme Bitcoin returns induce attention distraction for stock market investors. Investors tend to allocate more attention to the market-level information instead of firm-specific information during the extreme Bitcoin return days, which results in an increased change in stock return co-movement. This finding stretches our understanding of the role of attention allocation in connecting the cryptocurrency market and the traditional stock market. Second, we find that the COVID-19 pandemic is an amplifying factor for attention distraction. The outbreak of the COVID-19 positively moderates the "category-learning" behavior when extreme Bitcoin returns occur. This finding suggests that multiple attention shocks tend to reinforce each other to deepen investors' rational inattention to the stock market, which provides valuable insights into the behavioral research on stock investors.

Tables

Table 1. The Average Extreme Bitcoin and S&P500 Returns in the Bottom [Top] Rolling Percentiles

Year	Rolling 1% Percentile		Rolling 5% Percentile		Rolling 10% Percentile		Rolling 15% Percentile		Rolling 20% Percentile	
	Bitcoin	S&P500	Bitcoin	S&P500	Bitcoin	S&P500	Bitcoin	S&P500	Bitcoin	S&P500
2015	-12.71% [8.90%]	-2.38% [3.17%]	-9.31% [6.72%]	-2.04% [1.89%]	-7.04% [5.30%]	-1.61% [1.62%]	-5.85% [4.88%]	-1.38% [1.42%]	-4.87% [4.20%]	-1.26% [1.27%]
2016	-8.98% [6.05%]	-2.72% [1.99%]	-5.94% [5.53%]	-1.99% [1.74%]	-4.19% [4.68%]	-1.49% [1.54%]	-3.28% [3.99%]	-1.18% [1.34%]	-2.83% [3.16%]	-0.99% [1.18%]
2017	-13.01% [14.72%]	-1.30% [1.10%]	-8.93% [10.09%]	-0.82% [0.96%]	-7.37% [7.85%]	-0.66% [0.82%]	-6.54% [6.96%]	-0.53% [0.76%]	-5.56% [6.32%]	-0.47% [0.67%]
2018	-14.37% [10.11%]	-2.41% [2.03%]	-9.38% [7.43%]	-2.08% [1.88%]	-8.35% [6.65%]	-1.77% [1.51%]	-7.40% [5.92%]	-1.50% [1.36%]	-6.27% [5.00%]	-1.36% [1.21%]
2019	-11.35% [12.95%]	-2.49% [2.49%]	-7.27% [8.34%]	-1.82% [1.74%]	-5.94% [6.49%]	-1.38% [1.38%]	-5.16% [5.25%]	-1.21% [1.24%]	-4.22% [4.51%]	-1.03% [1.15%]
2020	-14.40% [9.92%]	-4.74% [4.32%]	-8.50% [7.34%]	-3.93% [3.48%]	-6.00% [5.88%]	-3.24% [3.26%]	-4.71% [5.07%]	-2.76% [2.67%]	-4.27% [4.54%]	-2.39% [2.48%]
2021	-12.18% [11.04%]	-2.19% [1.92%]	-8.10% [9.29%]	-1.68% [1.58%]	-7.17% [7.78%]	-1.34% [1.40%]	-6.07% [6.53%]	-1.15% [1.31%]	-5.36% [5.94%]	-0.99% [1.23%]

Notes: This table presents the average extreme daily returns for Bitcoin and S&P500 index. For each row, the numbers above represent the average daily returns above the 1-p^{th} percentiles in the previous 100-day rolling window, and the numbers bracketed below represent the average daily returns below the p^{th} percentiles in the previous 100-day rolling window.

Table 2. Abnormal Google Search Volume around Extreme Bitcoin Returns

	Bottom/Top 1% Percentile	Bottom/Top 5% Percentile	Bottom/Top 10% Percentile	Bottom/Top 15% Percentile	Bottom/Top 20% Percentile
	(1)	(2)	(3)	(4)	(5)
[-3,-1] days pre-event window	1.77% (0.59%)***	0.78% (0.30%)***	0.40% (0.24%)	0.09% (0.21%)	0.00% (0.20%)
During extreme returns	5.76% (0.69%)***	3.07% (0.37%)***	1.99% (0.30%)***	1.23% (0.26%)***	0.83% (0.24%)***
[+1, +3] days post-event window	4.37% (0.67%)***	1.95% (0.37%)***	1.10% (0.29%)***	0.56% (0.26%)**	0.26% (0.24%)
[-3,-1] days pre-and during extreme returns	2.91% (0.59%)***	1.53% (0.31%)***	0.96% (0.25%)***	0.52% (0.22%)**	0.37% (0.20%)*
During & [+1, +3] days post-event window	4.77% (0.65%)***	2.31% (0.36%)***	1.39% (0.29%)***	0.82% (0.25%)***	0.51% (0.23%)**
[-3, +3] around extreme returns	3.53% (0.57%)***	1.71% (0.31%)***	1.01% (0.25%)***	0.54% (0.22%)**	0.34% (0.20%)*

Notes: This table presents the summary of abnormal Google search volume (*AbGSVI*), the mean, and the standard error in the parenthesis, around extreme Bitcoin daily returns. ***/**/* indicate statistical significance at the 0.01/0.05/0.10 levels, respectively.

Table 3. Change in The Stock Return Co-movement on Extreme Bitcoin Return Days and Non-Extreme Return Days

Extreme Bitcoin Return Rolling Percentile Range:	Mean Correlation Coefficients			Mean Adjusted R^2			Significant $Cor_i^{BtcExtRet} - Cor_i^{NonBtcExtRet}$		
	Non-Extreme Returns	Extreme Returns	Diff (SE)	Non-Extreme Returns	Extreme Returns	Diff (SE)	# Negative Change (%)	# Positive Change (%)	# Insignificant Change (%)
	(1)	(2)	(3) = (2)-(1)	(4)	(5)	(6) = (5)-(4)	(7)	(8)	(9)
top/bottom 1%	0.351	0.482	0.131 *** (0.003)	0.156	0.281	0.124 *** (0.002)	300 (6.44%)	2722 (58.47%)	1633 (35.08%)
top/bottom 5%	0.314	0.417	0.102 *** (0.002)	0.132	0.223	0.091 *** (0.001)	335 (6.18%)	3328 (61.35%)	1762 (32.48%)
top/bottom 10%	0.297	0.392	0.094 *** (0.002)	0.119	0.200	0.080 *** (0.001)	357 (6.24%)	3581 (62.64%)	1779 (31.12%)
top/bottom 15%	0.295	0.372	0.077 *** (0.002)	0.118	0.181	0.063 *** (0.001)	389 (6.70%)	3418 (58.90%)	1996 (34.40%)
top/bottom 20%	0.295	0.357	0.063 *** (0.001)	0.119	0.167	0.048 *** (0.001)	474 (8.14%)	3201 (54.96%)	2149 (36.90%)

Notes: This table presents the stock return co-movement on extreme and non-extreme Bitcoin return days and their differences. Columns (1)-(3) report the change in stock return co-movement, measured by the Pearson correlation coefficient. Columns (4)-(6) report the change in stock return co-movement, measured by the adjusted R^2. Columns (7)-(9) report the distribution of the stocks by significant negative, significant positive, and insignificant change based on the Person correlation coefficient measure at the 95% confidence level. ***/**/* indicate statistical significance at the 0.01/0.05/0.10 levels, respectively.

Table 4. Decomposition of the Change in Correlation Coefficients on Extreme Bitcoin Return Days

Extreme Bitcoin Return Rolling Percentile Range:	Mean			Median		
	A (p-value)	*B (p-value)*	*C (p-value)*	*A (p-value)*	*B (p-value)*	*C (p-value)*
	(1)	(2)	(3)	(4)	(5)	(6)
top/bottom 1%	236.84% *** (0.00)	-32.63% *** (0.00)	-110.31% *** (0.00)	212.68% *** (0.00)	-18.21% *** (0.00)	-96.64% *** (0.00)
top/bottom 5%	227.92% *** (0.00)	-28.91% *** (0.00)	-100.00% *** (0.00)	199.07% *** (0.00)	-17.01% *** (0.00)	-87.30% *** (0.00)
top/bottom 10%	229.95% *** (0.00)	-28.90% *** (0.00)	-98.26% *** (0.00)	199.91% *** (0.00)	-19.02% *** (0.00)	-82.41% *** (0.00)
top/bottom 15%	226.40% *** (0.00)	-28.81% *** (0.00)	-94.36% *** (0.00)	190.05% *** (0.00)	-14.91% *** (0.00)	-76.15% *** (0.00)
top/bottom 20%	190.20% *** (0.00)	-9.61% *** (0.00)	-80.98% *** (0.00)	172.77% *** (0.00)	-7.86% *** (0.00)	-69.08% *** (0.00)

Notes: This table reports the results of cross-sectional tests on the percentage contributions from three components of the change in correlation coefficients on extreme Bitcoin return days, as discussed in Eq(3). *A*, *B*, and *C* denote covariance, stock return volatility, and market return volatility components, respectively. Each component is independently winsorized at 97.5% and 2.5% percentiles. Columns (1)-(3) report the cross-sectional t-test on the mean, and Columns (4)-(6) report the Wilcoxon signed-rank test on the median. ***/**/* indicate statistical significance at the 0.01/0.05/0.10 levels, respectively.

Table 5. Change in The Stock Return Co-movement During on Extreme Bitcoin Return Days: Pre-COVID Period and COVID Period

Panel A: Pre-COVID Sub-sample (2015-2019)

Extreme Bitcoin Return Rolling Percentile Range:	Mean Correlation Coefficients			Mean Adjusted R2			# Significant Co-Movement Change		
	Non-Extreme Returns	Extreme Returns	Diff (SE)	Non-Extreme Returns	Extreme Returns	Diff (SE)	# Negative Change (%)	# Positive Change (%)	# Insignificant Change (%)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
top/bottom 1%	0.319	0.430	0.111 *** (0.002)	0.133	0.216	0.083 *** (0.002)	180 (4.52%)	1936 (48.66%)	1863 (46.82%)
top/bottom 5%	0.296	0.374	0.078 *** (0.002)	0.119	0.175	0.056 *** (0.001)	241 (5.27%)	2261 (49.40%)	2075 (45.34%)
top/bottom 10%	0.293	0.346	0.054 *** (0.002)	0.117	0.155	0.037 *** (0.001)	363 (7.73%)	2052 (43.70%)	2281 (48.57%)
top/bottom 15%	0.287	0.340	0.053 *** (0.001)	0.114	0.150	0.037 *** (0.001)	341 (7.17%)	2206 (46.41%)	2206 (46.41%)
top/bottom 20%	0.294	0.317	0.023 *** (0.001)	0.119	0.133	0.014 *** (0.001)	637 (13.35%)	1530 (32.06%)	2606 (54.60%)

Panel B: COVID Sub-sample (2020-2021)

Extreme Bitcoin Return Rolling Percentile Range:	Mean Correlation Coefficients			Mean Adjusted R2			# Significant Co-Movement Change		
	Non-Extreme Returns	Extreme Returns	Diff (SE)	Non-Extreme Returns	Extreme Returns	Diff (SE)	# Negative Change (%)	# Positive Change (%)	# Insignificant Change (%)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
top/bottom 1%	0.433	0.551	0.118 *** (0.004)	0.226	0.360	0.134 *** (0.003)	280 (7.68%)	1822 (50.00%)	1542 (42.32%)
top/bottom 5%	0.384	0.494	0.110 (0.003)***	0.187	0.305	0.118 *** (0.002)	298 (6.96%)	2427 (56.68%)	1557 (36.36%)
top/bottom 10%	0.354	0.471	0.117 *** (0.002)	0.162	0.279	0.117 *** (0.002)	265 (5.86%)	2807 (62.03%)	1453 (32.11%)
top/bottom 15%	0.357	0.442	0.085 *** (0.002)	0.165	0.249	0.084 *** (0.002)	342 (7.46%)	2493 (54.37%)	1750 (38.17%)
top/bottom 20%	0.347	0.437	0.091 *** (0.002)	0.158	0.242	0.084 *** (0.002)	278 (6.03%)	2603 (56.45%)	1730 (37.52%)

Notes: This table presents the results of stock return co-movement analysis (in Table 3) on the pre-COVID (2015-2019) and the COVID (2020-2021) sub-samples in Panel A and Panel B, respectively. For additional details, see the notes in Table 3.

Table 6. Regression of Stock Return Co-movement on Extreme Bitcoin Return Indicator and COVID-19 Indicator

Panel A: Correlation Coefficients as Target Variable

Extreme Bitcoin Returns Rolling Percentile Range:	**5% Percentile** (1)	**10% Percentile** (2)	**15% Percentile** (3)	**20% Percentile** (4)
D^{COVID}	0.128***	0.107***	0.115***	0.096***
	(0.009)	(0.010)	(0.009)	(0.009)
$D^{BtcExtRet}$	0.078***	0.054***	0.053***	0.023***
	(0.004)	(0.004)	(0.004)	(0.003)
$D^{BtcExtRet} \times D^{COVID}$	0.032***	0.063***	0.0319***	0.068***
	(0.006)	(0.008)	(0.0048)	(0.005)
Fixed-Effects:	Stock	Stock	Stock	Stock
Clustered S.E.:	FF49 Ind	FF49 Ind	FF49 Ind	FF49 Ind
N	17,718	18,442	18,676	18,768
F-Statistics	83.702	97.298	103.540	105.660
R^2	0.8367	0.8563	0.8638	0.8661
Adj.R^2	0.7646	0.7917	0.8023	0.8058

Panel B: Adj R^2 as Target Variable

Extreme Bitcoin Return Rolling Percentile Range:	**5% Percentile** (1)	**10% Percentile** (2)	**15% Percentile** (3)	**20% Percentile** (4)
D^{COVID}	0.101***	0.082***	0.086***	0.073***
	(0.006)	(0.006)	(0.005)	(0.005)
$D^{BtcExtRet}$	0.057***	0.037***	0.037***	0.014***
	(0.007)	(0.005)	(0.005)	(0.003)
$D^{BtcExtRet} \times D^{COVID}$	0.061***	0.080***	0.047***	0.070***
	(0.007)	(0.011)	(0.008)	(0.010)
Fixed-Effects:	Stock	Stock	Stock	Stock
Clustered S.E.:	FF49 Ind	FF49 Ind	FF49 Ind	FF49 Ind
N	17,718	18,442	18,676	18,768
F-Statistics	73.8840	83.5310	91.8850	95.4130
R^2	0.8190	0.8364	0.8491	0.8538
Adj.R^2	0.7390	0.7629	0.7810	0.7880

Notes: This table presents the regression results in Eq(*4*). The regression analysis controls for stock fixed-effects and uses standard error clustering by Fama-French 49 industry classification. Numbers in the parenthesis are standard errors of the coefficients. ***/**/* indicate statistical significance at the 0.01/0.05/0.10 levels, respectively. Results remain qualitatively unchanged when using 2-year pre-COVID sub-samples (2018-2019).

Table 7 Decomposition of the Change in Correlation Coefficients on Extreme Bitcoin Return Days: Pre-COVID and COVID Periods

Panel A: Pre-COVID Sub-sample (2015-2019)

Extreme Bitcoin Return Rolling Percentile Range:	Mean			Median		
	A (p-value)	*B (p-value)*	*C (p-value)*	*A (p-value)*	*B (p-value)*	*C (p-value)*
top/bottom 1%	153.67% *** (0.00)	15.62% *** (0.00)	-68.81% *** (0.00)	130.76% *** (0.00)	11.81% *** (0.00)	-51.18% *** (0.00)
top/bottom 5%	142.23% *** (0.00)	8.60% *** (0.00)	-47.60% *** (0.00)	125.57% *** (0.00)	7.28% *** (0.00)	-42.87% *** (0.00)
top/bottom 10%	131.59% *** (0.00)	3.52% (0.18)	-32.12% *** (0.00)	112.66% *** (0.00)	8.33% *** (0.00)	-27.67% *** (0.00)
top/bottom 15%	142.43% *** (0.00)	-1.76% (0.52)	-41.77% *** (0.00)	126.38% *** (0.00)	3.40% ** (0.01)	-38.77% *** (0.00)
top/bottom 20%	100.00% *** (0.00)	22.12% *** (0.00)	-18.11% *** (0.00)	100.38% *** (0.00)	13.22% *** (0.00)	-15.40% *** (0.00)

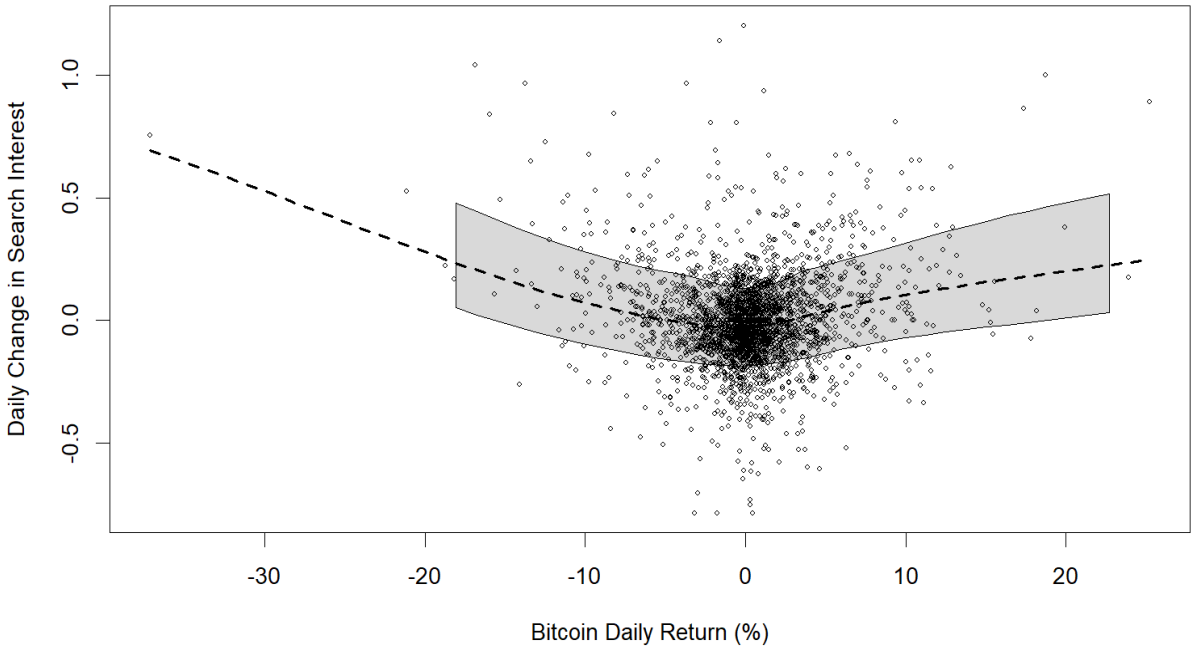
Panel B: COVID Sub-sample (2020-2021)

Extreme Bitcoin Return Rolling Percentile Range:	Mean			Median		
	A (p-value)	*B (p-value)*	*C (p-value)*	*A (p-value)*	*B (p-value)*	*C (p-value)*
top/bottom 1%	256.49% *** (0.00)	-46.90% *** (0.00)	-104.04% *** (0.00)	257.55% *** (0.00)	-37.27% *** (0.00)	-121.85% *** (0.00)
top/bottom 5%	326.79% *** (0.00)	-69.80% *** (0.00)	-155.98% *** (0.00)	267.92% *** (0.00)	-41.53% *** (0.00)	-128.19% *** (0.00)
top/bottom 10%	301.11% *** (0.00)	-57.77% *** (0.00)	-141.51% *** (0.00)	268.18% *** (0.00)	-37.18% *** (0.00)	-126.38% *** (0.00)
top/bottom 15%	264.00% *** (0.00)	-39.79% *** (0.00)	-119.39% *** (0.00)	221.78% *** (0.00)	-18.58% *** (0.00)	-101.66% *** (0.00)
top/bottom 20%	238.15% *** (0.00)	-25.38% *** (0.00)	-112.76% *** (0.00)	214.68% *** (0.00)	-16.04% *** (0.00)	-100.18% *** (0.00)

Notes: This table presents the results of decomposition analysis (in Table 4) on the pre-COVID (2015-2019) and the COVID (2020-2021) sub-samples in Panel A and Panel B, respectively. For additional details, see the notes in Table 4. Results remain qualitatively unchanged when using 2-year pre-COVID sub-samples (2018-2019).

Figures

Figure 1. Search Interest of "Bitcoin" and Bitcoin Returns



References

Anton, M., & Polk, C. (2014). Connected stocks. *The journal of finance, 69*(3), 1099-1127.

Barberis, N., Shleifer, A., & Wurgler, J. (2005). Comovement. *Journal of Financial Economics, 75*(2), 283-317.

Baur, D. G., Hong, K., & Lee, A. D. (2018). Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money, 54*, 177-189.

Choi, H. (2021). Investor attention and bitcoin liquidity: Evidence from bitcoin tweets. *Finance Research Letters, 39*, 101555.

Conlon, T., & McGee, R. (2020). Safe haven or risky hazard? Bitcoin during the COVID-19 bear market. *Finance Research Letters, 35*, 101607.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The journal of finance, 66*(5), 1461-1499.

Dastgir, S., Demir, E., Downing, G., Gozgor, G., & Lau, C. K. M. (2019). The causal relationship between Bitcoin attention and Bitcoin returns: Evidence from the Copula-based Granger causality test. *Finance Research Letters, 28*, 160-164.

Durnev, A., Morck, R., Yeung, B., & Zarowin, P. (2003). Does greater firm‐specific return variation mean more or less informed stock pricing? *Journal of Accounting Research, 41*(5), 797-836.

Guo, X., Lu, F., & Wei, Y. (2021). Capture the contagion network of bitcoin–Evidence from pre and mid COVID-19. *Research in International Business and Finance, 58*, 101484.

Hu, Y., Li, X., Goodell, J. W., & Shen, D. (2021). Investor attention shocks and stock co-movement: substitution or reinforcement? *International Review of Financial Analysis, 73*, 101617.

Hu, Y., Li, X., & Shen, D. (2020). Attention allocation and international stock return comovement: evidence from the Bitcoin market. *Research in International Business and Finance, 54*, 101286.

Huang, S., Huang, Y., & Lin, T.-C. (2019). Attention allocation and return co-movement: Evidence from repeated natural experiments. *Journal of Financial Economics, 132*(2), 369-383.

Huang, Y., Duan, K., & Mishra, T. (2021). Is Bitcoin really more than a diversifier? A pre-and post-COVID-19 analysis. *Finance Research Letters, 43*, 102016.

Jennrich, R. I. (1970). An asymptotic $\chi 2$ test for the equality of two correlation matrices. *Journal of the American Statistical Association, 65*(330), 904-912.

Klein, T., Thu, H. P., & Walther, T. (2018). Bitcoin is not the New Gold–A comparison of volatility, correlation, and portfolio performance. *International Review of Financial Analysis, 59*, 105-116.

Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of Financial Economics, 58*(1-2), 215-260.

Nguyen, K. Q. (2021). The correlation between the stock market and Bitcoin during COVID-19 and other uncertainty periods. *Finance Research Letters*, 102284.

Núñez, J. A., Contreras-Valdez, M. I., & Franco-Ruiz, C. A. (2019). Statistical analysis of bitcoin during explosive behavior periods. *PloS one, 14*(3), e0213919.

Ozkan, O. (2021). Impact of COVID-19 on stock market efficiency: Evidence from developed countries. *Research in International Business and Finance, 58*, 101445.

Peng, L., & Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics, 80*(3), 563-602.

Philippas, D., Rjiba, H., Guesmi, K., & Goutte, S. (2019). Media attention and Bitcoin prices. *Finance Research Letters, 30*, 37-43.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics, 50*(3), 665-690.

Smales, L. A. (2020). Investor attention and the response of US stock market sectors to the COVID-19 crisis. *Review of Behavioral Finance*.

Smales, L. A. (2021). Investor attention and global market returns during the COVID-19 crisis. *International Review of Financial Analysis, 73*, 101616.

Urquhart, A. (2018). What causes the attention of Bitcoin? *Economics Letters, 166*, 40-44.

Vidal-Tomás, D. (2021). Transitions in the cryptocurrency market during the COVID-19 pandemic: A network analysis. *Finance Research Letters, 43*, 101981.

Zhang, D., Hu, M., & Ji, Q. (2020). Financial markets under the global pandemic of COVID-19. *Finance Research Letters, 36*, 101528.

Zhang, X., Lu, F., Tao, R., & Wang, S. (2021). The time-varying causal relationship between the Bitcoin market and internet attention. *Financial Innovation, 7*(1), 1-19.

Journal of Risk and Financial Management

MDPI

Article

Financial Crises and Business Cycle Implications for Islamic and non-Islamic bank lending in Indonesia

Samar Issa

Department of Business Administration, Saint Peter's University, Jersey City, NJ 07306, USA; sissa@saintpeters.edu

6/21/2022

Abstract: This paper addresses if and how excess-debt as defined by Stein (2012a) can be considered as an early warning signal for banks and takes an additional dimension by comparing the excess leverage between Islamic and conventional banks in Indonesia before, during, and after the Global Financial Crisis (GFC). To do so, this research develops an empirical model of banks' capital structure and optimal- and excess-debt, and conducts an empirical analysis on the overleveraging of eleven conventional and Islamic banks in Indonesia. Results show that all banks became vulnerable to the GFC in 2007–2009 while credit build-up, i.e., overleveraging, started in 2005. However, for most of the banks in the sample, Islamic banks performed better and leveraged less prior to the GFC, which made them more resilient to the crisis.

Keywords: banking instability; banking sector; Islamic banking; credit flows; financial crisis; excess debt; early warning signals

1. Introduction

Over the past two decades, excessive capital borrowing (leveraging) has risen significantly for the financial sector globally. Following the 2008 Global Financial Crisis (GFC), there has been a rise in academic and policy research on the problems of debt sustainability. The COVID-19 pandemic has reignited the urgency of the topic. While more solid empirical assessments in relation to the pandemic are yet to come, this paper asks the question whether Islamic banks, which are argued to be safer than conventional banks for having their product structured as asset-backed financing, have maintained a more sustainable debt level pre- and post-GFC. Many studies investigated issues related to asset-price channels through which the banking system's instability is triggered. Some of these important academic contributions include Brunnermeier and Sannikov (2014), Mittnik and Semmler (2012, 2013), Stein (2012a, 2012b), and Gross et al. (2017).

To undertake the empirical study on those issues in Indonesia, this paper uses a measure of overleveraging as defined by Stein (2012). Essentially, if borrowing exceeds debt capacity, then this can be called "excess leveraging." Debt capacity is measured as sustainable or optimal-debt. In other words, the measure of overleveraging is defined as the difference between actual- and optimal-debt. The model presented in this paper follows Stein (2012b) and Issa (2020), and focuses on the solution of the dynamic version of the Stein (2012b) model that allows us to use time-series data on banks in order to calculate the excess debt of eleven banks in Indonesia.

In this paper, overleveraging refers to lending booms. It is well known that lending booms may precede banking system instability as they imply increased risk-taking in the financial system. This has the potential to result in financial turmoil if the economy is hit by a negative adverse shock in asset prices as occurred during the GFC of 2008.

Banks have continued to leverage after 2012 and there are expectations of another financial crisis affecting the banking sector following the COVID-19 crisis; we have already seen the effect on the financial market. However, it is important to differentiate between the nature of the external COVID-19 shock and the 2008 crisis. The former is a purely exogenous shock while the latter is a result of a vulnerable state of the financial system. COVID-19 provides a shock that is not connected to business or financial cycles and, as such, behavior can be analyzed empirically with less concern for the usual endogeneity issues. It must be noted that the government should control leverage in order to prevent a future crisis of any type (see what happened in COVID-19). The COVID-19 pandemic delivered a structurally different impact, but still layering upon subdued endogenous fragilities. For the main endeavor of this paper, this paper proceeds with the core Stein model as laid out in Equations (1) and (2), leaving the COVID-19 crisis and the extension opportunities for future projects.

How did the overleveraging of financial intermediaries occur in Indonesia? After the bust of the technology boom in 2000, not only asset prices rose globally but financial intermediaries increased their loan supply to the real estate sector. Since the 2008 GFC and subsequent quantitative easing by developed countries, large capital inflows into emerging market economies (EMEs) were induced. Many of the financial and real trends that occurred after 2001 are documented in Semmler and Bernard (2012) and Stein (2011). The extent to which these countries are affected depends on each country's financial system and macroeconomic conditions. For instance, Indonesia is a bank-based economy characterized by shallow financial markets, which posits the economy at a higher volatility and instability. Moreover, global factors have played a role in adjusting the country's monetary policy through a number of channels such as the interest rate, exchange rate, asset prices, bank deposit and lending, and debt issuance as well as risk-taking and excessive leverage.

This paper addresses if and how excess-debt as defined by Stein (2012a) can be considered as an early warning signal for banks and takes an additional dimension by comparing the excess leverage between Islamic and conventional banks in Indonesia before, during, and after the GFC. Results showed that all banks became vulnerable to the GFC in 2007–2009, whereas credit build-up, i.e., overleveraging, started in 2005. However, for most of the banks in the sample, Islamic banks performed better and leveraged less during the crisis, which meant these banks were affected less by the crisis and also made them more resilient to a financial crisis. Nevertheless, Islamic banks still in some cases relied on leveraging and were subject to the second-round effect of the global crisis around the years of 2011–2013, when higher excess-debt values were found for Islamic banks. That could also be due to the fact the financial crisis hit the real economy after 2009 because most Islamic bank investments and contracts were backed by real estate and property as collateral. The results show that the challenges in Indonesia became even more complex in 2013 as they coincided with the domestic problem of inflation pressure from an increase in domestic fuel prices and a widening current account deficit due to the decline in commodity prices together with booms in lending and house prices.

This paper is organized as follows. Following the introduction and rationale of this study in Section 1, Section 2 gives the literature review about Islamic banking in Indonesia. A theoretical model of optimal leveraging is developed in Section 3. Section 4 presents empirical estimation results with some interpretations and Section 5 concludes the paper. The appendices provide the technical background and calculations of excess-debt.

2. Literature Review

The expansion of Islamic banking has motivated several researchers to assess Islamic banks' performance. A recent study by Marlina et al. (2021) aimed to analyze the growing trend of available financial research regarding Islamic banking. The authors used bibliometric mapping (Borner et al., 2003) and found that as of January 2021, there were 500 published scientific research papers, with journal articles being the great majority. They found that the most common authors were Sukmana,

Ascarya, and Ismal. The keywords most often used were "Indonesian" and "Islamic." The newest trend topics circle around governance, sharia issue, and the social role of Islamic banking in Indonesia. As Islamic banking is growing at a fast rate, the new research contains many beneficial findings that can be implemented to fix and improve issues within the financial industry and, in return, improve people's lives.

An International Monetary Fund (IMF) paper by Bitar et al. (2017) studied the effect of Basel Core Principle (BCP) compliance on the stability of conventional and Islamic banks. The authors used a sample of 761 banks in 19 countries for the years 1999–2013. Their findings suggest that the BCP index is positively associated with the stability of conventional banks, whereas the effect is less noticeable for Islamic banks.

Another IMF paper by Shabsigh et al. (2017) showed that the complexity of Islamic banks exposed them to additional risks that should be addressed by supervisory authorities and central banks. Islamic banks' hybrid products exposed them not only to additional returns but also to added risks that made them volatile. The authors argued that there was a need to establish a clear framework for troubled Islamic banks to ensure the stability of the financial system.

A number of recent theoretical studies have embarked to explore the Islamic banks in Indonesia specifically. For instance, Kusuma et al. (2016) investigated the relationship of bank performance with its corporate governance. The authors selected a sample of eleven Islamic banks for the years 2010–2014 and performed a regression analysis on the data extracted from balance sheets of banks. Findings showed that the efficiency level of corporate governance of Indonesian Islamic banks improved during the period in context. In addition, corporate governance efficiency was significantly correlated to Islamic banks' performance.

A recent paper by Abduh et al. (2021) took a macroeconomic approach and tried to find the correlation between Islamic banks' short-term and long-term performance, and Indonesian economic growth. The methodology in this research used quarterly data from 2003 to 2010, extracted from International Financial Statistics of IMF and from the monthly Islamic Banking Statistics Report of the Bank of Indonesia. The variables representing the Islamic banks were total financing and the indicators representing the Indonesian economy were the Gross Domestic Product and Gross Fixed Capital Formation representing investments. The authors found a relationship between Islamic banking and Indonesian economic growth. However, the surprising aspect was that the relationship was bi-directional, meaning that Islamic banks promoted economic growth while economic growth itself also stimulated Islamic banking. Thus, both supply-leading and demand-following relationships were taking place at the same time. In order to use this information, the author suggested that the Indonesian government should incorporate Islamic banking more in Indonesia as well as review the current regulations because they have been shown to have no benefit with the corresponding growth.

Furthermore, Cakranegara (2020) attempted to analyze the severity of the damage caused by Covid-19 in the financial sector by comparing the Covid-19 crisis with the 1998 crisis. The purpose of the study was to identify similarities between the two crises that would allow regulators and managers to draw lessons from the past and apply it to the current crisis. The author claimed that Indonesia maintained stability through its system, consisting of "several layers" such as the debt selection system, the loan distribution system, the minimum capital owned by a bank, the rule to use asset insurance, the banks' health monitoring system, and government intervention when everything else failed (Rosas, 2006). Therefore, Indonesian banks diversified risk through different industries, periods, and products in order to lower their systemic risk. Nevertheless, when the Covid-19 pandemic hit, every industry was hit and, thus, all banks were affected. The author stated that another approach is needed in order to prevent the same fall from happening in a future crisis. This paper found that Indonesian banks nowadays are far more resilient and that people also find them more reliable than they were in 1998. The lower effect of the crisis was due to the improved capital structure of the banks as well as the more lenient actions taken by the government. The author

believed that strict government measures in 1998 contributed to the banking recession and stated that public trust and credit are key contributors to economic growth.

Maghfuriyah et al. (2019) similarly pointed out that short-term and long-term market structures affected banking performance in Islamic banking. To achieve that, the authors used the Structure Conduct Performance analysis by implementing an error correction model. They utilized monthly time-series data from 2015 to 2018. The data were acquired from the Islamic banking monthly report, the Islamic banking statistics published by the Financial Service Authority, and the financial statement of each Islamic bank in Indonesia. The market structure variables used in the study were the concentration ratio, the Herfindahl–Hirschman Index, and the market share. Concentration ratio was defined as "the percentage of the overall industrial output generated by the largest companies" and the Hirschman Index was used to obtain an accurate description of the analysis concentration ratio. The error correction model showed that in the short term, market structure had an insignificant effect on banking performance growth. However, in the long term, the opposite occurred; market structure had a great effect on banking performance due to certain determinants. For example, if the market share of deposits increased, then the Islamic banking performance would decrease in the long term but increase in the short term. Moreover, if the concentration ratio increased, then the Islamic banks performance would decrease both in the short and long term. On the other hand, if the market concentration ratio decreased, then the banking performance would increase. With such findings, the authors hoped that regulators and firms could make better informed decisions to allow the Islamic banks to flourish both in the short and long term.

The paper by Nugroho et al. (2018) showed how segmentation affected the quality of Islamic banks based on non-performing financing ratio/non-performing loan (NPF/NPL). The research used a quantitative method with multiple regression tests and demonstrated that retail segmentation held more weight than whole segmentation did in influencing the quality of the banks (i.e., 92.61% and 56.05%, correspondingly). Moreover, the study stated that Islamic banks faced greater challenges than conventional ones. In 2017, starters only held 10% of the market share while conventional banks held 90%. Not to mention that some key financial indicators, such as low-quality assets, showcased the financing channeled by Islamic banks. Based on NPF Gross, "problematic financing in (the Islamic) banking industry grew by 341.85% over eight years from 2008 to 2015. The authors concluded that given their findings, it was critical that for the retail segment to find a "special treatment" to maintain high quality of assets and Islamic bank performance. In other words, banks should reinforce and follow up with their sharia principles in contributing to social and environmental aspects to engage more customers with their services.

Another paper by Rodoni et al. (2017) focused on finding the efficiency rate and productivity of Islamic banks given their reliance on that to compete against conventional banks. The model used time-series secondary data extracted from sources such as bank reports, statements publications, and other relevant sources. The technique used by the authors to get the efficiency rates was the Data Envelopment Analysis and for productivity was the Malmquist Index (MI). The study refers to data from 2009 to 2013. The findings showed that most of the Islamic banks had inefficient rates. The Indonesian Islamic banks ranged between 87% and 97% in efficiency rates and never reached 100%, which the authors attributed to external rather than internal factors. Moreover, the authors stated that during the last four years, there has been a trend of increasing productivity but that such was caused by managerial factors rather than technological ones. As for Malaysian Islamic banks, the situation was somewhat better. Their efficiency rates ranged from 92% to 95% even though in the last 5 years they have not reached 100% efficiency rate either. The percentage of inefficiency was

attributed to external factors as well. Lastly, there has been a fluctuating and a likely negative growth in productivity that is mostly due to technological issues.
Third, Pakistani Islamic banks have performed better compared to the Indonesian and the Malaysian banks. The efficiency rate varied between 99.3% and 100% during the last 5 years. As for productivity, their trend has been positive and increasing due to technological advancements and good management.

Mulyaningsih et al. (2011) analyzed the competitiveness of the Indonesian banking industry based on consolidation and concentration. The authors claimed that in 1990, banks experienced a structural change due to the deregulation policy in 1988, making it easier to enter the banking market. However, in 1995, the government started to tighten the regulations once again. In 1997 (post-1995 crisis), the Central Bank merged four state-owned banks and closed the operation of twenty-three others in order to achieve economies of scale and healthier systems. To increase stability in the banking industry, the Central Bank focused on increasing monitoring based on international standards and fostering good corporate governance. To assess the competitiveness of the banks, the authors used a regression model. The authors found that the structure of the Indonesian market was vulnerable, given that the market was concentrated in a few banks because large banks control a substantial share. Second, the consolidation actually led to a less concentrated market and better market share. Third, during the consolidation policy, banks were running under "monopolistic competition" for large banks, but after some time, competition started growing for all. The authors showed that a concentrated market led to less competition, which explained why larger banks in Indonesia who had higher concentration also had less competition.

Empirically, a paper by Asutay et al. (2007) analyzed the performance (profitability) of Bank Muamalat in Indonesia. The authors' methodology relied on a regression analysis to estimate the internal and external determinants in reference to its return on assets. The research used monthly bank-specific data extracted from balance sheets and income statements from 1996 to 2001 while macroeconomic indicators were derived from the Bank of Indonesia. The authors found that the bank's profit largely came from financial activities and not from service activities. Moreover, from 1996 to 2001, the bank's portfolio relied on too much short-term-based financing. A positive correlation between the ratio of total liabilities to total assets showed that Islamic banks have incentives to take on more risks. Nevertheless, a negative coefficient of this same ratio showed that Islamic banks tended to lose money when they offered more of the Mudarabah schemes. The authors' findings also matched with the hypothesis that there is a direct relationship between inflation and profitability measure.

3. Empirical Approach and Definition of Variables

3.1. Background of financial crisis and Islamic banking

According to many researchers, the GFC crisis was caused by excessive financial obligations/mortgages of private households (i.e., bubbles in the mortgage market defined as unsustainable debt/income ratios). Stein (2012b) argued that although debt problems may have originated in either the public or private sectors in different nations, the result was still declining asset values and the mechanisms at work resulted in a contagion effect either from the United States to Europe and/or from one European nation to another, depending on the debtor-to-debtee relationship under examination. Of course, in each scenario, Stein made it clear that the primary source of the problem was not the presence of debt but excess debt within the country/countries under analysis.

Stein derived an optimal-debt ratio and built on it to identify an early warning signal (EWS) of a debt crisis, which is defined as the "excess debt of households" (actual-debt ratio less than the optimal-debt ratio). As the excess-debt level rises, the probability of a debt crisis increases. It has been shown that rising house prices since the late 1990s led to above-average capital gains for households, thereby increasing owner equity. The supply of mortgages increased and, consequently, financial obligations as a percentage of disposable income increased for private households. At the same time, the quality of loans declined (subprime mortgages). Of course, this process was not sustainable. As capital gains[1] dropped below the interest rate, debtors could no longer service their debts and foreclosures led to a collapse in the value of financial derivatives.

Before delving into the theoretical model and empirical analysis presented in this paper, it is important to briefly provide background information regarding Islamic banks' mortgage-loan operations. As mentioned, their operations do not involve interest rates; rather, they have a proprietary program called the LARIBA (interest-free) model, which uses equity-participation systems or profit/loss sharing. Islamic banks offer two types of mortgage loans. First, profit sharing (*murabaha*) is offered whereby the bank does not loan money to the buyer to purchase the home or other property; rather, the bank buys the home itself and then resells it to the buyer at a profit. The buyer typically pays a fairly large down payment and the price at which the bank buys the property is disclosed to the end buyer. The second method is decreasing rent (*ijara*). The bank purchases the home and resells it to the buyer; however, unlike the first method, in this case, the home remains in the bank's name until the total price is paid. The buyer takes up residence immediately and makes payments to the bank on the purchase price, but in addition to the payments, the buyer also pays a fair market rent. This method is preferred in countries such as the United Kingdom to avoid the double payment of taxes (Khan, 2010).

Dr. Abdul-Rahman, Chairman and CEO of the Bank of Whittier, said during an interview that they had few nonperforming mortgages in 2008 and that this was due to their riba-free (RF) discipline[2]. He explained that the bank tracks home prices in US dollars relative to more stable commodities, such as gold, silver, wheat, or rice, thereby allowing for the discovery of real estate bubbles. He added that on the basis of the used strategy, the bank had an early warning signal of a macro-bubble forming in the real estate sector. In addition, they matched the property to the market, meaning that they researched the rental value of a similar property in the same neighborhood. He stated that the inputs of their model are the amount to be invested (financed), the number of years of financing, and the rent. The unknown is the rate of return on investment[3]. Their internal business model allowed them to survive the recent financial crisis.

3.2. Theoretical Model

This paper introduces a model of optimal leverage that helps us define overleveraging. The model sketched here is a low-dimensional stochastic variant of a model of banking leveraging, and it follows Stein (2008, 2012a) and Brunnermeier and Sannikov (2014).

This paper focuses on the Stein (2008, 2012) model. However, several recent literatures develop models to estimate excess-debt, for both private and public debt, and many of them stem from the same (2008, 2012) Stein model.

For instance, Issa and Gevorkyan (2022) develop an empirical model of corporate capital structure, optimal debt, and overleveraging, covering approximately the period from 2000 to 2018, across the following six industries: technology, financial, pharmaceutical, auto, airline, and energy. The authors estimate optimal and excess debt for each of the 89 firms and find that corporate excess debt has largely been moving up, spiking around the GFC and continuing into recovery more

[1] Meaning increases in housing prices.

[2] See "The Art of Riba-Free (RF) Islamic Banking and Finance", published by Wiley (first e. 2010 and second e. 2014) by Dr. Yahia Abdul-Rahman.

[3] If the rate is much lower than the accepted return in the market, say 4–5%, they do not invest/finance. If the rate is high, then they finance, but reduce the rent such that the monthly payment would be competitive with that of other banks in the market.

recently. The trend is consistent with an increase in the actual debt, with varying average excess debt ratios by sector.

An IMF paper by Eberhardt and Presbitero (2018) develops an empirical model using regression analysis to predict banking crises. The authors employ a sample of 60 low-income countries for the time span 1981–2015 and show that credit growth cannot be considered a main driving force of economic distress or financial crisis, as most literature states. The reason is that credit growth is mediated through capital inflow, which is also fueled by a booming financial market. Therefore, banking crises are potentially driven by commodity price changes according to the authors.

Another paper by Nyambuu and Bernard (2015) uses the Stein (2008) model to calculate optimal debt for developing countries. In contrast to this paper, which calculates the optimal debt at the micro-level for the baking sector, the authors apply the model at a macro-level. In both cases, the optimal debt ratios are defined as the threshold beyond which it becomes risky to borrow, or in other words, the distance-from-default indicator variable. While both papers calculate optimal debt, each methodology applies different metrics, since this paper addresses banking default while Nyambuu and Bernard (2015) report sovereign default. Nyambuu and Bernard (2015) show that rising ratios of external debt can make a country vulnerable to shocks and increases the risk of default. Similarly, this paper shows that an increase in excess debt of a certain bank will increase its risk to default on its debt.

Finally, Gevorkyan and Semmler (2016) examine the U.S. shale energy sector, and their model is consistent with Stein's optimal debt argument as well. The framework is rooted in the theory that the boom trend in borrowing leads to overleveraging and risks of insolvency. Employing a nonlinear model predictive control (NMPC) analysis, the authors present two mutually exclusive scenarios: oligopoly or extensive competition. The empirical analysis of actual and optimal debt based on a vector error correction model (VECM) suggests that firms with larger market capitalization in the industry are more resilient to oil price volatility than medium and small-sized firms.

Overall, this model is very similar to those of Brunnermeier and Sannikov (2014) and Stein (2012a). Both models have leveraging and payouts as choice variables, and net worth as a state variable. Moreover, both models are stochastic. Similar to this study, Brunnermeier and Sannikov (2014) specifically focused on the banking sector; however, the setting is more general compared to the one used in this paper. There are households that save, and financial experts representing financial intermediaries that invest in capital assets owned by households and financial intermediaries. Both have different discount rates. I focus solely on the behavior of financial intermediaries.

In this model, preferences used in the objective function and Brownian motions as state variables similar to both studies. The Stein (2012a) model, assuming certain restrictions, uses log utility and allows to exactly compute excess leveraging. Capital return is also stochastic due to capital gains, and the interest rate is stochastic as well, similar to my model and in contrast with that of Brunnermeier and Sannikov (2014); where only the capital return is stochastic, and the interest rate is taken as constant. Both Brunnermeier and Sannikov (2014) and Stein (2012a) employed a continuous time version, but the problem in this paper is formulated as a discrete time variant with a discounted instantaneous payout and an optimal leveraging function.

Moreover, a shock to asset prices creates a vicious cycle through the balance sheets of banks. In other words, risk taking and excessive borrowing occur when asset prices are volatile. In this model, the unconstrained growth of capital assets through excessive borrowing, facilitated by the lack of regulations imposed on financial intermediaries, is considered the main cause for banking-sector instability[4]. On the other hand, large payouts with no "skin in the game" affect banks' risk-taking behaviors, equity development, and leveraging. The higher the payout is, the more leveraged the bank becomes, which increases the aggregate risk and risk premia for all. In summary, the increased risk spreads, and risk premia, especially at a time when defaults begin, expose banks to vulnerabilities and financial stress triggered by security-price movements.

[4] This source of instability is also discussed by Brunnermeier and Sannikov (2014).

To derive an optimal-debt ratio, Stein used stochastic optimal control (SOC). A hypothetical investor selects an optimal-debt ratio, $f(t)$, to maximize the expectation of a concave function of net worth, $X(t)$, where T is the terminal date. The model assumes that the optimal-debt/net-worth ratio significantly depends on the stochastic process concerning the capital-gain variable. The expected growth of net worth is also maximal when the debt ratio is at the optimal level.

Optimal leverage is given by:

$$f^*(t) = [(r - i) + \beta - \alpha y(t) - \frac{\left(\frac{1}{2}\right)\left(\sigma_p^2 - \sigma_i \sigma_p \rho\right)}{\sigma^2}] \quad (1)$$

such that

$$Risk = \sigma^2 = \sigma_i + \sigma_p - (2\rho_{ip}\sigma_i\sigma_p), \quad (2)$$

where r is the bank's capital gain/loss; i is the credit cost of banks; β is the productivity of capital; $y(t)$ is the deviation of capital gain from its trend; σ^2 is the variance; and ρ represents the negative-correlation coefficient between interest rate and capital gain[5]. Through the presented model, Stein could determine excess debt and an early warning signal of a potential crisis. As mentioned, it is this mechanism that played a role in the decreasing net worth of individuals, households, and institutions in the United States, and that was amplified by the increased leverage and pricing volatility of complex securities.

To measure the excess leveraging of banks, the introduced and defined Stein model was followed with a focus on the solution of the dynamic version of the model, which allowed for using time-series data on banks. One difference from Stein is that, in this case, each bank's productivity of capital was not assumed to be deterministic or constant as in the Stein model; rather, it was calculated for the years of 2000–2018.

The optimal-debt level was calculated for the years 2000 until 2018; thus, excess debt, which is the measure of overleveraging in this paper, was estimated. To calculate the banks' optimal-debt ratios, data on the banks' capital gain/loss, market interest rates, and the productivity of capital were collected. Using these variables, the risk and return components of the model were then calculated. Using the abovementioned variables, the optimal-and actual-debt ratios were calculated for a sample of eleven banks, four Islamic and seven conventional banks[6]. The full calculations are presented in 11 tables[7] with 18 columns each. Data were derived from the banks' annual reports, balance sheets, and financial statements, in addition to some data received directly from bank managers. The calculations are summarized below.

Column 1 consists of capital gain/loss that represents the return in percentage to the investors of the bank from capital appreciation or loss in a particular year. This capital gain/loss is calculated by dividing the change in each bank's stock-market cap by the beginning market cap at each period. The market caps were Hodrick–Prescott (HP)-filtered to eliminate the effects of daily stock-market swings. HP Filter is a data-smoothing technique frequently applied on time series data to remove short-term fluctuations associated with the business cycle.

Column 2 represents the market interest rate. The 10-year treasury yield was used to represent the market interest rate, and is therefore presented in percentage[8].

In Column 3, beta (β) represents the productivity of capital. The beta is calculated as the bank's annual gross revenue divided by total capital. The total capital here is calculated as the shareholder equity plus half of both short-term and total long-term debt[9]. To determine shareholder equity, I

[5] The full mathematics of the Stein model (2008, 2012a) are provided in Appendix A.

[6] The list of banks' names and types is provided in Appendix B.

[7] The full calculations are presented in Tables A2–A13 in Appendix C.

[8] Variables in the first two columns, capital gain/loss and interest rate, form the uncertainty of the model. The two variables are stochastic in the model and can move in different directions.

[9] The reason that total capital is calculated this way is because capital investments in a company are made up of equity capital and debt financing; hence, a company has two types of stakeholders: equity and debt holders.

obtained the annual value of each bank's shareholder equity from the balance sheet. Short-term debt comprises all the banks' current liabilities that are usually due within 12 months. Long-term debts, on the other hand, are calculated as the combinations of long-term liabilities and other liabilities in banks' balance sheets. These are basically all bank liabilities due in more than a year's time. Therefore, each bank's productivity of capital is calculated for the years of 2000–2016, and not constant as in Stein (2012b).

Columns 4 through 9 are the risk elements in the model (Stein 2012a). Column 4 represents beta variance that calculated as the difference between each year's beta from the mean beta for the years 2000 and 2018, representing the deviation of each period's beta from the mean. Column 5 is also a component of the risk element. This is calculated as one-half of the square of the capital-gain variable. Column 6 is the statistical correlation between interest-rate and capital-gain variables over the period from 2000 to 2018[10]. Columns 7 and 8 are the variance for the interest-rate and capital-gain variables, respectively. Each period's variance is calculated as the deviation of that period's value from the mean. Therefore, interest-rate variance is the difference between each year's 10-year treasury yield and the mean interest rate from 2000 to 2018. Similarly, capital-gain variance is the difference between each year's capital gain/loss and the mean capital gain from 2000 to 2018.

Column 9, which is the product of the correlation between the stochastic variables (interest rate and capital gain), and interest-rate and capital-gain variance, represents an additional component of the risk element. It is calculated as the product of the correlation factor of the stochastic variables (Column 6), interest-rate variance (Column 7), and capital-gain variance (Column 8)[11]. Columns 10–12 are used to determine the risk-investors bear when they decide to hold equity in the bank, and this is a key issue for the investors' decision making.

Columns 10 and 11 represent the standard deviations of the interest rate and capital gains, respectively. Therefore, Column 10 is the standard deviation of values in Column 2, while Column 11 is the standard deviation of values in Column 1. Here, standard deviations are constant over the periods, as in the Stein model.

Column 12, on the other hand, is calculated as twice the value of variances of the two stochastic variables and their correlation. This is, therefore, calculated as 2 multiplied by Column 9.

Column 13 is, hence, the risk of an investor holding the equity of the bank at each time period as in Equation (A2). The risk is calculated using Columns 10–12. The risk is calculated by adding the standard deviations of the interest rate (Column 10) plus the standard deviation of capital gain (Column 11) minus the risk component in Column 12.

In the model, the optimal-debt ratio maximizes the difference between net return and risk term. Therefore, only if the net return exceeds the risk premium does the optimal-debt ratio become positive. The optimal-debt ratio, therefore, is not a constant, as Stein also noted (2012a), but rather varies directly with net return and risk.

In Column 14, I then calculated, using all the above-mentioned variables, Stein optimal-debt ratio, f*(t). Debt ratios were normalized to remove the effects of seasonality. Therefore, normalized f*(t) measures the deviation of the optimal-debt ratio away from the mean. Negative values in Column 14 represent lower optimal-debt ratios away from the mean ratio during the applicable periods. The components of the optimal-debt ratio are, therefore, primarily the capital gains for equity holders of the bank's stock, the market interest rate, and the risk term. The optimal-debt ratio maximizes the difference between mean return and risk term. The formula for calculating optimal-debt ratio using the above column numbers is: ((Column 1 − Column 2) + Column 3 − Column 4 − Column 5 + Column 9)/Column 1. This reiterates what was mentioned above that optimal-debt ratio is positive only if the net return is greater than the risk premium, and this can intuitively be seen.

In Column 15, I calculated normalized optimal-debt ratios using Column 14, the mean and standard deviation of the optimal-debt values[12]. In addition to calculating the optimal-debt ratio, I calculated the banks' actual-debt ratio in order to calculate the excess-debt ratio. The actual-debt ratio

[10] In the calculation, this correlation is a constant value over the period.

[11] The optimal-debt ratio is positive only if the net return is greater than the risk premium.

[12] They were normalized such that each variable had a mean of zero.

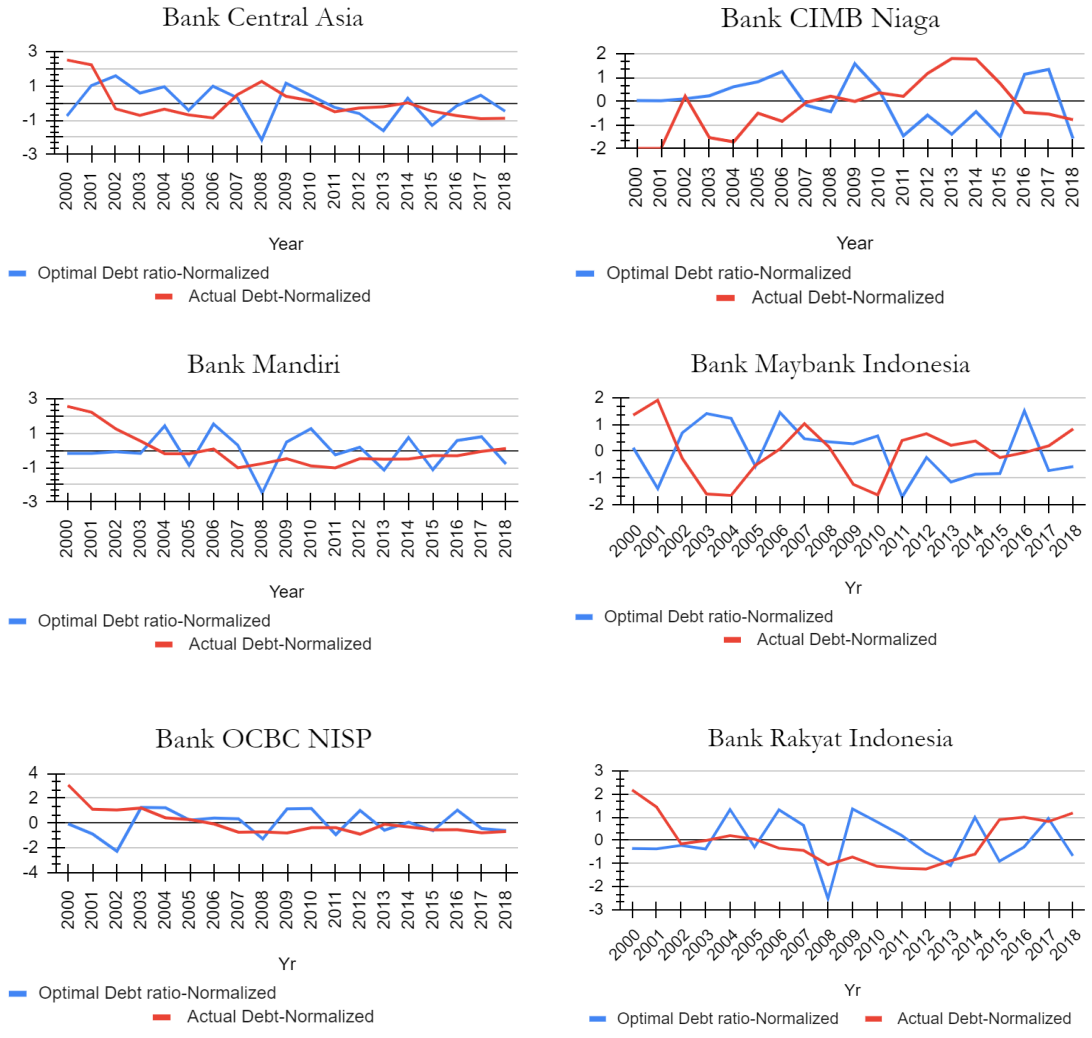
of the banks was equal to long-term debt divided by total assets, which are given in the banks' annual reports as well. Actual-debt ratios are also normalized in the same way as optimal-debt ratios are, and are presented in Column 16. After optimal- and actual-debt ratios are calculated as discussed above, excess debt is calculated in the last columns as normalized actual minus optimal debt. The graphs of the two ratios, namely, actual- and optimal-debt ratios, are presented in Figures 1–5 below, followed by empirical analysis.

4. Empirical Analysis

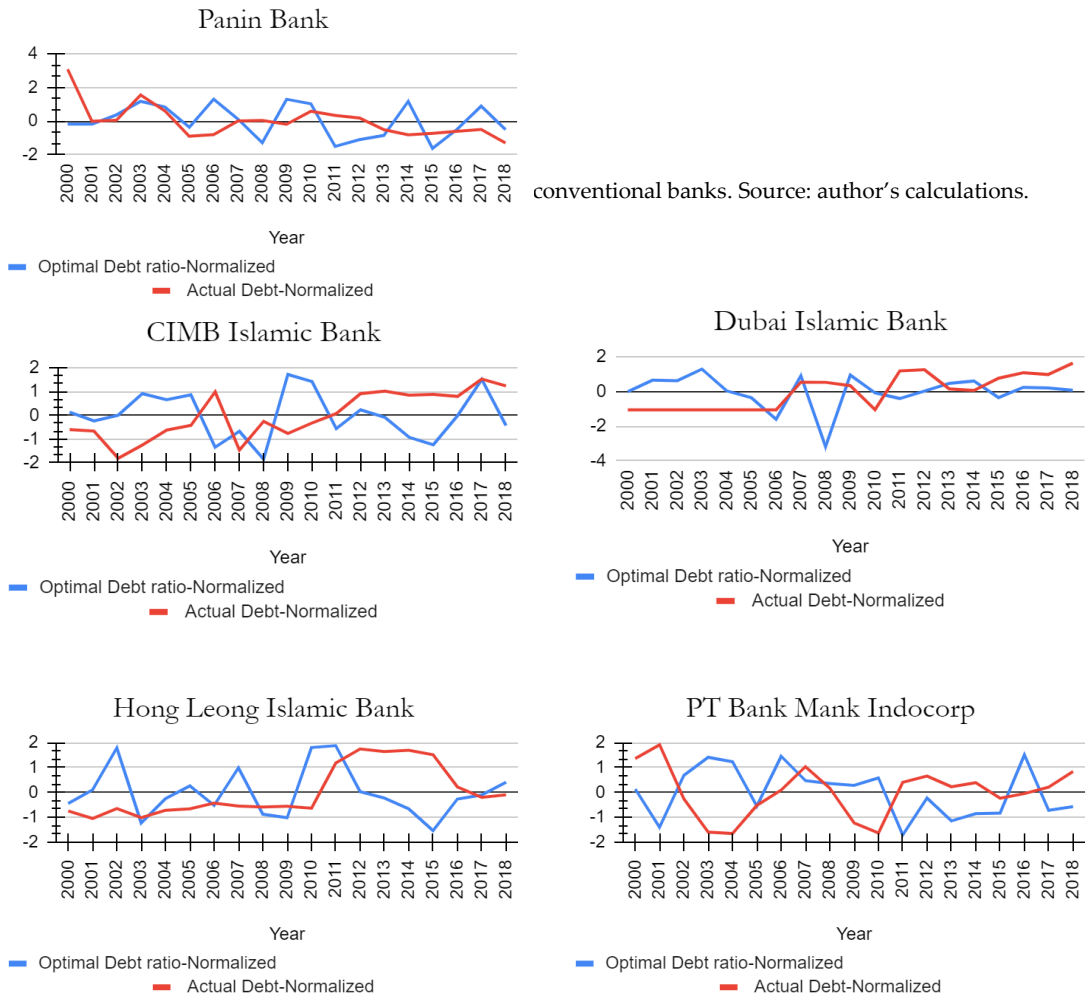
4.1. Graphical Results and Analysis: Actual- vs. Optimal-debt

Next, applying the methodology presented in the previous section, optimal leverage was calculated for a sample of eleven banks in Indonesia. For the banks under study, this analysis was performed using the banks' total long-term debts and total assets. As noted, total long-term debt represents a bank's total debt with a maturity date of more than one year from the balance sheet date. Total asset represents the value of a bank's total assets[13].

The vertical axes of Figure 1 represent the debt ratios while horizontal axes represent the years. First, Figure 1 shows the optimal debt against actual-debt ratios for the seven conventional banks under study in Indonesia. Similarly, Figure 2 shows optimal-debt against the actual-debt ratios for the four Islamic banks under study in Indonesia.



[13] Such as cash, accounts receivable, investments in other firms, properties, and intangible assets.



conventional banks. Source: author's calculations.

Figure 2. Optimal- versus actual-debt ratio for Islamic banks. Source: author's calculations.

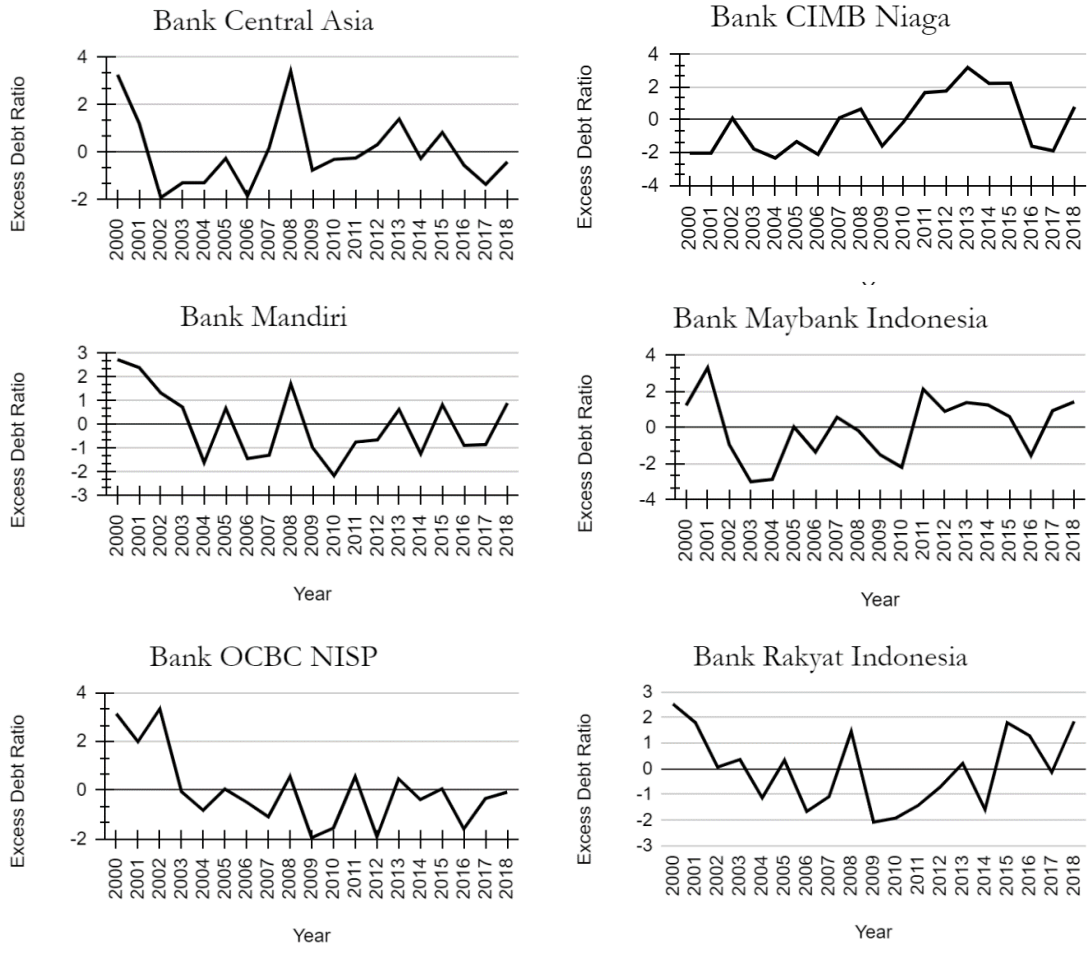
The graphs show the optimal- against the actual-debt ratio for the eleven banks under study for the period 2000–2018. The optimal-debt ratios for most of the banks exhibited similar trends although the trends were more pronounced for conventional than for Islamic banks. For a number of years preceding the 2007–2009 financial crisis, these banks had high optimal-debt ratios. For most of the banks, about a year or two prior to 2007, optimal-debt ratios began to drop, and the decrease was severe in most cases. Overall, optimal-debt for most Islamic banks was not as high as optimal-debt for conventional banks in this context. On the other hand, the drop in optimal-debt right before the crisis for both types of banks exhibited the same behavior. Another interesting observation is that optimal-debt for Islamic banks was more stable during the years immediately following the crisis when compared with that of conventional banks, with the exception of a couple of conventional banks. Moreover, the trend of actual-debt exceeding optimal-debt clearly reversed post-crisis for most banks in the sample.

Each of the banks' actual-debt ratio was also calculated. As Figures 1 and 2 show, most of the conventional banks exhibited an increase in actual-debt ratio over the years prior to the crisis, with a

smoother and less severe increase in the debt ratio for the Islamic banks. More specifically, the Bank of Central Asia, Mandiri, NISP, and Panin Bank have kept the levels of optimal- and actual-debt ratios close. As for the rest of the traditional banks, actual-debt is apart from the optimal-debt, either below the optimal ratio or above it. Overall, it seems that periods of low leverage follow excess leverage periods as banks try to keep their debt below the optimal ratio to balance out the different peaks of high debt, way above the optimal course. In general, there are five peaks of overleveraging periods for traditional banks, whereas there are around six periods of stable debt levels over time. It is noticed that traditional banks' ratios had gone up toward the absolute value of 3 while Islamic banks reached a maximum ratio at below the absolute value of 2. Thus, Islamic banking seems less volatile, with an actual-debt level orbiting around the respective optimal level for each bank, unlike traditional banks in many instances. Overall, we notice three periods of overleveraging that Islamic banks have experienced.

4.2. Graphical Results and Analysis: Excess-debt

For further analysis, the deviation of the banks' actual-debt ratios from the mean over the period 2000–2018 was calculated as an alternative proxy for excess-debt ratio. Excess-debt ratios were calculated for the eleven above-mentioned banks. For this calculation, each bank's total assets and long-term debts were used to calculate debt ratios as previously noted. Total assets represent the bank's total balance sheet assets at the end of the period. Long-term debts are the balance of each bank's debts that are more than one year due at the end of each period. On the basis of the presented calculation method, the excess-debt of each bank was calculated[14] and graphs of the excess-debt ratios are exhibited in Figures 3 and 4.



[14] See attached Appendix C for detailed calculations of excess debt in Tables A2–A13.

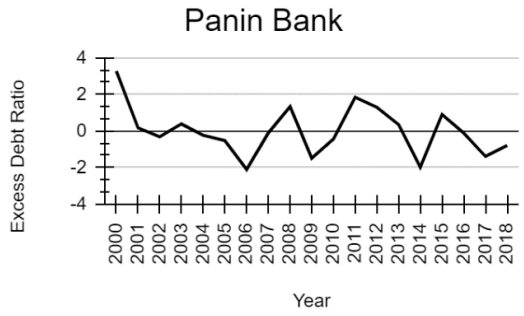


Figure 3. Excess-debt ratios for conventional banks. Source: author's calculations.

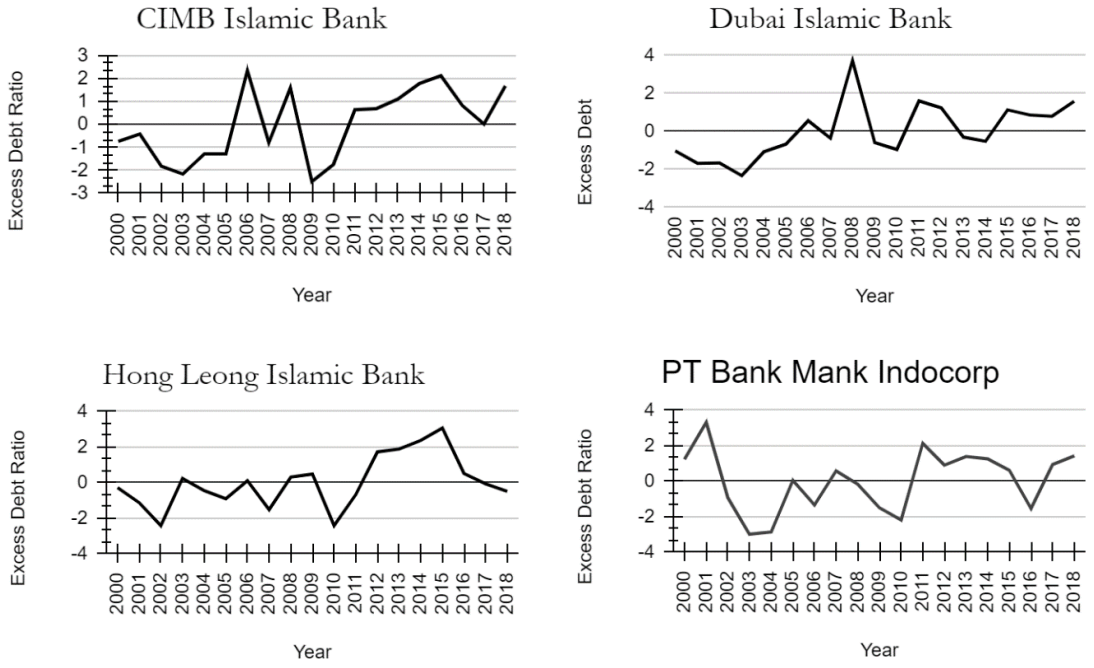


Figure 4. Excess-debt ratio for Islamic banks. Source: author's calculations.

In the graphs in Figures 3 and 4, vertical axes are the deviation of debt ratios, namely, excess-debt ratio, and the horizontal axes represent years. Graphs show that all banks in this study exhibited a similar movement in debt ratio for most of the period with excess-debt between 2007 and 2009 during the financial crisis.

For conventional banks, we observe a decline in excess-debt between the years 2000 and 2003. However, this reversed and peaked from 2005 until 2008 when excess-debt for all banks started to increase at a high rate. Right after the GFC, we see a sharp decline until 2010, where we can observe a rising trend once again during the period 2010–2015; however, the trend was not as steep as was observed in 2006–2008. After 2015, there was a period of recovery that continued until 2017. During this period, all conventional banks had sustainable debt levels but started to take risk and excessively leverage in 2018 and onwards. This can clearly be seen in the graphs above with the line of excess-debt rising for all traditional banks.

Generally, the Islamic banks followed a similar trend, with low levels of debt from 2000 to 2003. Then, from 2004 to 2006, some banks kept a constant debt while others exhibited an increase. From 2007 to 2008, there was a rise on the excess-debt but not as high as in the traditional banks as we can see from the graphs. All banks in Indonesia, including Islamic banks, had excess-debt in the years preceding and during the GFC, with excess-debt higher for Islamic banks after the GFC. This occurred because Indonesian Islamic banks also highly relied on leverage, which made them vulnerable to the second-round effect of the global crisis (Hasan and Dridi, 2010).

Most of the banks exhibited a rise in excess-debt between 2010 and 2012, overlapping a weak economic growth due to the global economy, with the sovereign-debt crisis in Europe and policy uncertainties in the United States impacting investments and companies. At times of low economic growth, companies begin to face liquidity issues and, therefore, request more loans, which increase the leverage of banks.

Table 1 and below showcases the highest excess-debt for conventional and Islamic banks, respectively, during the full period under study as well as the average high excess-debt and Table 2 presents the same for the conventional banks.

#	TRADITIONAL BANKS	HIGH EXCESS-DEBT YEAR(S)	DEBT RATIO	AVRG EXCESS RATIO
1	Bank Central Asia	2000	3.26	3.33
		2008	3.41	
2	Bank CIMB Niaga	2013	2.19	2.19
3	Bank Mandiri	2000	2.73	2.27
		2001	2.39	
		2008	1.70	
4	Bank Maybank Indonesia	2001	3.31	2.71
		2011	2.12	
5	Bank OCBC NISP	2000	3.13	3.23
		2002	3.33	
6	Bank Rakyat Indonesia	2000	2.52	1.89
		2001	1.80	
		2008	1.46	
		2015	1.80	
		2018	1.85	
7	Panin Bank	2000	3.27	2.14
		2008	1.33	
		2011	1.84	

Table 1: High Excess Leverage among Conventional Banks

#	ISLAMIC BANKS	HIGH EXCESS-DEBT YEAR(S)	DEBT RATIO	AVRG EXCESS RATIO
1	**CIMB Islamic Bank**	2006	2.33	2.02
		2008	1.59	
		2015	2.12	
2	**Dubai Islamic Bank**	2008	3.71	2.29

		2011	1.59	
		2018	1.56	
3	**Hong Leong Islamic Bank**	2012	1.71	2.36
		2015	3.01	
4	**PT Bank Mank Indocorp**	2001	3.31	2.71
		2011	2.12	

Table 2: High Excess Leverage among Islamic Banks

Table 1 presents the conventional banks that have had the highest leverage ratios. For instance, although Bank Central Asia and Bank OCBC NISP only had two years of high leverage, their ratios had been extremely high compared to the other banks. The average ratio for Central Asia and OCBC NISP was 3.33 and 3.23, respectively. On the other hand, we can see that although Bank Rakyat Indonesia had high leverage for five years—considerably more frequent high leverage compared to its peers—its average excess ratio was relatively lower.

As we can see from Table 2, the two Islamic banks, namely, CIMB Islamic Bank and Dubai Bank, had the highest leverage. The high leverage continued for three years as opposed to other banks, which only have two years of high overleverage. However, we notice that PT Bank Mank Indocorp had the highest average excess leverage ratio (2.71) among the Islamic banks.

Tables 3 and 4 exhibit the lowest debt levels based on these years for conventional and Islamic banks in our sample, respectively.

#	TRADITIONAL BANKS	LOWEST DEBT YEAR(S)	DEBT RATIO	AVRG DEBT RATIO
1	Bank Central Asia	2002	-1.92	-1.61
		2004	-1.30	
		2006	-1.85	
		2017	-1.36	
2	Bank CIMB Niaga	2000	-2.02	-1.91
		2001	-2.02	
		2003	-1.77	
		2004	-2.32	
		2006	-2.10	
		2009	-1.59	
		2016	-1.60	
		2017	-1.88	
3	Bank Mandiri	2004	-1.61	-1.63
		2006	-1.45	
		2007	-1.30	
		2010	-2.16	
4	Bank Maybank Indonesia	2003	-3.01	-2.09
		2004	-2.89	
		2006	-1.36	
		2009	-1.52	
		2010	-2.21	

		2016	-1.56	
5	Bank OCBC NISP	2009	-1.94	-1.74
		2010	-1.55	
		2012	-1.90	
		2016	-1.58	
6	Bank Rakyat Indonesia	2006	-1.66	-1.73
		2009	-2.08	
		2010	-1.92	
		2011	-1.43	
		2014	-1.58	
7	Panin Bank	2006	-2.10	-1.86
		2009	-1.49	
		2014	-1.98	

Table 3: Low Excess Leverage among Conventional Banks

#	ISLAMIC BANKS	LOW DEBT YEAR(S)	DEBT RATIO	AVRG DEBT RATIO
1	CIMB Islamic Bank	2003	-2.16	-2.16
		2002	-1.82	
		2009	-2.49	
2	Dubai Islamic Bank	2001	-1.71	-1.91
		2002	-1.68	
		2003	-2.35	
3	Hong Leong Islamic Bank	2002	-2.43	-2.14
		2007	-1.54	
		2010	-2.44	
4	PT Bank Mank Indocorp	2003	-3.01	-2.24
		2004	-2.89	
		2009	-1.52	
		2010	-2.21	
		2016	-1.56	

Table 4: Low Excess Leverage among Islamic Banks

When looking at Table 3, we see that Bank CIMB Niaga had the longest time of having low-leverage. For eight years, its ratio was low, below optimal-debt. However, the lowest average ratio was the Bank of Maybank Indonesia with a ratio of −2.09.

From the above, we can see that the years 2000, 2001, 2011, and 2008 have the highest overleverage ratio for both traditional and Islamic banks. The worst year, 2008, had the highest overleverage with six banks having an excess ratio greater than 1.3. (See Appendix C for detailed calculations.)

However, it can be noted that most Islamic banks did better than their counterparts given that they had a lower frequency of high overleverage.

On the other hand, we see the years 2009 and 2010 reflect the highest frequency of banks with low leverage, where the majority of banks were conventional. An explanation can be that after periods of high leverage before 2008, banks' sustainability to lend decreased, which forced the banks to lower their debt levels. Notice how in 2008 there were zero banks with low leverage.

As we can see from Table 4, PT Bank Mank Indocorp had the longest time of having a low leverage compared to the rest of the banks, although the bank exhibited the highest average overleveraging ratio during 2001 and 2011. We can observe how after those years of credit booms, a period of low leverage and credit crunch occurred in order to balance out the overleveraging and bring the bank to stability once again. Empirically, a long period of deleveraging nearly always follows a major financial crisis and these episodes have been painful.

In conclusion, we clearly see that conventional banks showed high excess-debt before and during the GFC, whereas Islamic banks maintained a low excess-debt level during and after the crisis. This is due to their lower leverage and higher solvency, which allowed them to meet relatively stronger demand for credit; however, for both types of banks, an increase in excess-debt a couple of years preceding the crisis and/or during the crisis was observed although the level of increase differed among banks. Overall, Islamic banks leveraged less than conventional banks. To maintain high profits, conventional banks relied on loans, especially loans with high risks. This made the banks even more vulnerable due to the real estate boom. In addition, banks increased lending for corporate takeovers and leveraged buyouts, which are highly leveraged transaction loans.

5. Conclusions

With respect to our main endeavor, empirical analysis showed that leverage is lower in general at Islamic banks compared with their conventional counterparts in Indonesia. This study presented a model that helped identify the early warning signs of banking crises on the basis of the presence of excess-debt or what is called "overleveraging." Here, a measure of overleveraging was defined as the difference of actual and sustainable debt. Furthermore, an empirical study was conducted for eleven banks in Indonesia that examined the vulnerability of these banks, and the credit and output contractions that could subsequently arise when hit by a crisis.

Results showed that all banks in the sample exhibited high excess-debt levels preceding the GFC. It was argued that excess-debt, rather than actual-debt, could serve as an early indicator of crises because the presence of excess-debt is actually the reason why banks collapse. Although banks have different debt ratios, a significant rise in excess-debt ratios often precedes and/or overlaps a banking crisis. Although Islamic banks exhibited a lower level of leverage, their operations were yet characterized by a high degree of financial risk. The reason is that Islamic banks undertake risky operations in order to be able to generate a return that can attract customers when they cannot be guaranteed a return on deposits. What might make Islamic banks unique and attractive in the coming years is the short-term funding and noninterest-earning assets that might be in high need and are considered the main aspects of Islamic banks' profits. The results of this paper show that Islamic and conventional banks in Indonesia had a consistent behavior with similar banks in Bahrain, Kuwait, Malaysia, the United States, and the United Kingdom (see Issa, 2020).

As seen above, the excess leverage taken by the banking sector in Indonesia poses serious threats. The unprecedented policy response to the COVID-19 pandemic has helped prevent a meltdown and maintain the flow of credit to the economy. However, the outlook remains highly uncertain and vulnerabilities are rising, representing a potential early warning signal. The performance evaluation of Islamic banks and the stability of these types of banks are especially important today. In the aftermath of all that is currently happening after the Covid-19 crisis, we can probably anticipate a rise in non-Western conventional banking/leveraging activities. It would be interesting for future research to look into a concrete indication through data; however, it would seem logical that as conventional funds are leaving emerging markets, these economies have to rely on their specific advantages to draw additional funds—here, the focus on Islamic banking would be ratified. A better understanding of these policy questions requires specific knowledge about the riskiness of the banks and the leverage level undertaken by Islamic banks as well as their profitability.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

Acknowledgment: The author thanks Tabata Patino for data collection.

Appendix A

Appendix A.1. Mathematical Derivation of Stein's Optimal-debt

Here, Stein shows how the optimal-debt ratio is derived in the logarithm case.
The stochastic differential equation is (A1)

$$dX(t) = X(t)\left[\left(1 + f(t)\right)\left(\frac{dP(t)}{P(t)} + \beta(t)dt\right) - i(t)f(t) - cdt\right] \quad \text{(A1)}$$

where the Debt Ratio is $f(t) = \frac{L(t)}{X(t)} = \frac{Debt}{Net\ Worth}$; Capital Gain or Loss is $\frac{dP(t)}{P(t)}$; Productivity of Capital is $\beta(t) = \frac{Income}{Assets}$; Interest Rate is $i(t)$; $\left(1 + f(t)\right) = \frac{Assets}{Net\ Worth}$; Ratio of Consumption is: $c = \frac{(Consumption\ or\ Dividends)}{Net\ Worth}$ and is taken as given.

Let the price evolve as:

$$dP(t) = P(t)\left(\alpha(t)dt + \sigma_p dw_p(t)\right), \quad \text{(A2)}$$

where $\alpha(t)$ represents the asset's drift component and the interest rate is represented by the sum of i and a Brownian Motion term as follows:

$$i(t) = idt + \sigma_i dw_i(t). \quad \text{(A3)}$$

Substitute (A2) and (A3) in (A1) and derive (A4):

$$dX(t) = X(t)\left[\left(1 + f(t)\right)\left(\left(\alpha(t)dt + \sigma_p dw_p(t)\right) + \beta(t)dt\right) - i(t)f(t)dt - cdt\right] \quad \text{(A4)}$$

$$dX(t) = X(t)\left(\mathrm{M}f(t)\right)dt + X(t)\beta f(t)$$

$$\mathrm{M}f(t) = \left[\left(1 + f(t)\right)\left((\alpha(t)dt) + \beta(t)dt\right) - i(t)f(t)dt - cdt\right]$$

$$\beta(t) = [\left(1 + f(t)\right)\sigma_p dw_p - \sigma_i f(t)dw_i(t)]$$

$$\beta^2 f(t) = [(1 + f(t)^2){\sigma^2}_p\, dt + f(t)^2{\sigma^2}_i\, dt - 2f(t)\left(1 + f(t)\right)\sigma_i\sigma_p dw_p dw_i]$$

$$\text{Risk= R}f(t) = \left(\frac{1}{2}\right)[(1 + f(t)^2){\sigma^2}_p\, dt + f(t)^2{\sigma^2}_i\, dt - 2f(t)\left(1 + f(t)\right)\sigma_i\sigma_p].$$

$\mathrm{M}f(t)$ contains the deterministic terms and $\beta(t)$ contains the stochastic terms. To solve for $X(t)$, consider the change in $\ln X(t)$ in (A5). This is based upon the Ito equation of the stochastic calculus. A great virtue of using the logarithm criterion is that one does not need to use dynamic programming. The expectation of $dlnX(t)$is (A6):

$$dlnX(t) = \left(\frac{1}{X(t)}\right)dXt - (\frac{1}{2}X(t)^2)(dx(t)^2) \quad \text{(A5)}$$

$$E\left[d\left(lnX(t)\right)\right] = [\mathrm{M}f(t)] - \mathrm{R}[f(t)]dt. \quad \text{(A6)}$$

The correlation $\rho dt = E(dw_p dw_i)$ is negative, which increases risk: $(dt)^2 = 0, dwdt = 0$.
The optimal-debt ratio f^* maximizes the difference between the Mean and Risk:

$$f^* = argmax_f\left[M\left(f(t)\right) - R\left(f(t)\right)\right] = [\alpha(t) + \beta(t) - i] - [\left(\frac{(\sigma_p^2 - \rho\sigma_i\sigma_{p)}}{\sigma^2}\right)]$$

$$f^* = argmax_f\left[M\left(f(t)\right) - R\left(f(t)\right)\right] = f^*(t) \quad \text{(A7)}$$
$$= \{(r - i) + \beta - \alpha y(t) - (\frac{1/2(\sigma_p^2 - \rho_{ip}\sigma_i\sigma_{p)}}{\sigma^2})\}$$

s.t.

$$Risk = \sigma^2 = {\sigma_i}^2 + {\sigma_p}^2 - \left(2\rho_{ip}\sigma_i\sigma_p\right).$$

Appendix A.1.1. Model I

Model I assumes that the price $P(t)$ has a trend rt and a deviation $Y(t)$ from it (A8). The deviation $Y(t)$ follows an Ornstein–Uhlenbeck ergodic mean reverting process (A9). Coefficient α is positive and finite. The interest rate is the same as in model II:

$$P(t) = Pexp\left(rt + y(t)\right). \quad \text{(A8)}$$

The deviation from the trend is demonstrated through:

$$y(t) \; = \; ln\, P(t) - \; ln\, P - \; rt.$$

The mean reversion aspect characterized by a convergence of α is defined as:

$$dy(t) = -\alpha(t)dt + \sigma_p dw_p(t). \quad \text{(A9)}$$

In this model, Stein defines E(dw) = 0; $E(dw)^2 = dt$:

$$limy(t) \sim N\left(0, \frac{\sigma^2}{2\alpha}\right).$$

Stein constrains the solution such that $r \leq i$ and calls this the "No free lunch constraint." Therefore, using the stochastic calculus in model I is the first term in (A10):

$$dP(t) = \; P(t)\left(\alpha(t)dt + \; \sigma_p dw_p(t)\right)$$
$$dP(t)/P(t) = (r - \alpha y(t)) + \frac{1}{2}{\sigma_p}^2)dt + \sigma_p dw_p\,, \quad \text{(A10)}$$

where $\alpha(t)$ represents the asset's drift component and the interest rate is represented by the sum of i and a Brownian Motion term as follows:

$$i(t) = idt + \sigma_i dw_i(t)\,.$$

Substituting (A10) in (A7) and deriving (A11), the optimal-debt ratio in model I is as follows:

$$f^*(t) = \left[(r - i) + \beta - \alpha y(t) - \frac{\left(\frac{1}{2}\right)\left(\sigma_p^2 - \sigma_i\sigma_p\rho\right)}{\sigma^2}\right]. \quad \text{(A11)}$$

Consider $\beta(t)$ as deterministic.

Appendix A.1.2. Model II

In model II, the price equation is (A12). The drift is $\alpha(t)dt = \boldsymbol{\pi}$dt and the diffusion is $\sigma_p dw_p$:

$$dP(t)/P(t) = \boldsymbol{\pi}\text{dt} \; + \sigma_p dw_p \quad \text{(A12)}$$

The optimal-debt ratio $f^*(t)$ is (A13). Consider $\beta(t)$ as deterministic:

$$f^*(t) = [(\boldsymbol{\pi} + \; \beta(t) - i) - \frac{\left(\sigma_p^2 - \sigma_i\sigma_p\rho\right)}{\sigma^2}] \quad \text{(A13)}$$

s.t.

$$\sigma^2 = {\sigma_i}^2 + {\sigma_p}^2 - \left(2\rho_{ip}\sigma_i\sigma_p\right).$$

In terms of a maximization portfolio decision, we have:

$$\underset{\alpha_t}{\text{Max}}[\alpha_t\left(E(R_{t+1}) - R_{F,t+1}\right) - \tfrac{k}{2}\alpha_t^2\sigma_t^2]. \quad \text{(A14)}$$

Appendix B

Table A1. List of Banks

Bank Name	Type	Country
Central Asia Bank	Conventional Bank	Indonesia
CIMB Bank Niaga	Conventional Bank	Indonesia
Mandiri Bank	Conventional Bank	Indonesia
Maybank Indonesia Bank	Conventional Bank	Indonesia
OCBC NISP Bank	Conventional Bank	Indonesia
Bank Rakyat Indonesia	Conventional Bank	Indonesia
Panin Bank	Conventional Bank	Indonesia
CIMB Bank	Islamic Bank	Indonesia
Dubai Islamic Bank	Islamic Bank	Indonesia
Hong Leong Bank	Islamic Bank	Indonesia
PT Mank Indocorp Bank	Islamic Bank	Indonesia

Appendix C

Table A2. Calculations of Optimal-and Excess-debt: CIMB Islamic Bank

	Year	Capital Gains/ (losses), (r)	Interest Rate (i)	Beta (Productivity of Capital, β)	Beta variance (σy(t))	Half Square of capital gain variance	Correlation of interest and capital gain variables	Interest rate variance	Capital gain variance	Correlation and variances of interest and capital gain	Std. deviation of interest rate	Std. deviation of capital gain	2 x (correlation and variances of interest and capital gain)	Risk	Optimal debt ratio, $f^*(t)$	Normalized Optimal Debt ratio	Actual Debt ratio	Normalized Actual Debt ratio	Excess Debt
CIMB Islamic Bank	2000	0.0096	0.0258	0.3368	(0.020)	0.000	(0.36)	(0.0098)	0.3260	0.0011	0.0133	0.6503	0.0023	0.661	0.52	0.13	0.0494	(0.61)	(0.74)
	2001	-0.0965	0.0243	0.3065	(0.051)	0.005	(0.36)	(0.0113)	0.0636	0.0003	0.0133	0.6503	0.0005	0.663	0.35	(0.25)	0.0485	(0.67)	(0.42)
	2002	-0.0424	0.0209	0.2710	(0.086)	0.001	(0.36)	(0.0147)	1.2205	0.0064	0.0133	0.6503	0.0127	0.651	0.46	(0.00)	0.0296	(1.82)	(1.82)
	2003	0.2783	0.0188	0.3168	(0.040)	0.039	(0.36)	(0.0168)	-0.1798	(0.0011)	0.0133	0.6503	(0.0021)	0.666	0.87	0.91	0.0388	(1.26)	(2.16)
	2004	0.1899	0.0286	0.3015	(0.056)	0.018	(0.36)	(0.0070)	-0.0292	(0.0001)	0.0133	0.6503	(0.0001)	0.664	0.75	0.66	0.0491	(0.63)	(1.28)
	2005	0.2486	0.0191	0.3377	(0.020)	0.031	(0.36)	(0.0165)	0.2895	0.0017	0.0133	0.6503	0.0034	0.660	0.84	0.86	0.0524	(0.43)	(1.28)
	2006	2.3655	0.0197	0.3522	(0.005)	2.798	(0.36)	(0.0159)	-0.2307	(0.0013)	0.0133	0.6503	(0.0026)	0.666	(0.14)	(1.35)	0.0755	0.99	2.33
	2007	-0.1993	0.0339	0.4041	0.047	0.020	(0.36)	(0.0017)	-0.6142	(0.0004)	0.0133	0.6503	(0.0007)	0.664	0.16	(0.68)	0.0352	(1.47)	(0.80)
	2008	-0.4600	0.0373	0.4205	0.063	0.106	(0.36)	0.0017	0.8440	(0.0005)	0.0133	0.6503	(0.0010)	0.665	(0.37)	(1.85)	0.0551	(0.26)	1.59
	2009	1.1797	0.0252	0.4292	0.072	0.696	(0.36)	(0.0104)	0.0327	0.0001	0.0133	0.6503	0.0002	0.663	1.23	1.72	0.0468	(0.77)	(2.49)
	2010	0.5658	0.0374	0.4022	0.045	0.160	(0.36)	0.0018	-0.5646	0.0004	0.0133	0.6503	0.0007	0.663	1.10	1.42	0.0540	(0.33)	(1.74)
	2011	-0.1541	0.0476	0.4411	0.084	0.012	(0.36)	0.0120	1.1362	(0.0049)	0.0133	0.6503	(0.0097)	0.673	0.21	(0.57)	0.0605	0.07	0.64
	2012	0.0637	0.0442	0.3922	0.035	0.002	(0.36)	0.0086	0.3179	(0.0010)	0.0133	0.6503	(0.0020)	0.665	0.56	0.23	0.0743	0.91	0.68
	2013	-0.0338	0.0422	0.3907	0.033	0.001	(0.36)	0.0066	0.1408	(0.0003)	0.0133	0.6503	(0.0007)	0.664	0.42	(0.09)	0.0760	1.01	1.10
	2014	-0.2548	0.0415	0.3313	(0.026)	0.032	(0.36)	0.0059	-0.0714	0.0002	0.0133	0.6503	0.0003	0.663	0.04	(0.93)	0.0732	0.84	1.77
	2015	-0.3276	0.0405	0.3714	0.014	0.054	(0.36)	0.0049	0.2717	(0.0005)	0.0133	0.6503	(0.0010)	0.664	(0.10)	(1.25)	0.0737	0.87	2.12
	2016	-0.0089	0.0504	0.3416	(0.016)	0.000	(0.36)	0.0148	-0.1599	0.0008	0.0133	0.6503	0.0017	0.662	0.45	(0.02)	0.0724	0.79	0.81
	2017	0.6666	0.0516	0.3112	(0.046)	0.222	(0.36)	0.0160	-0.1599	0.0009	0.0133	0.6503	0.0018	0.662	1.13	1.51	0.0843	1.52	0.01
	2018	-0.1107	0.0666	0.3293	(0.028)	0.006	(0.36)	0.0310	-0.1599	0.0018	0.0133	0.6503	0.0035	0.660	0.27	(0.43)	0.0796	1.24	1.67

Table A3. Calculations of Optimal- and Excess-debt: Dubai Islamic Bank

	Year	Capital Gains/ (losses), (r)	Interest Rate (i)	Beta (Productivity of Capital, β)	Beta variance (αγ(t))	Half Square of capital gain variance	Correlation of interest and capital gain variables	Interest rate variance	Capital gain variance	Correlation and variances of interest and capital gain	Std. deviation of interest rate	Std. deviation of capital gain	2 x (correlation and variances of interest and capital gain)	Risk	Optimal debt ratio, $f^*(t)$	Normalized Optimal Debt ratio	Actual Debt ratio	Normalized Actual Debt ratio	Excess Debt
Dubai Islamic Bank	2000	0.0000	0.0258	0.6658	0.278	0.000	(0.27)	(0.0098)	0.3260	0.0009	0.0133	0.7859	0.0017	0.797	0.46	(0.02)	0.0000	(1.06)	(1.04)
	2001	0.2875	0.0243	0.6203	0.232	0.041	(0.27)	(0.0113)	0.0636	0.0002	0.0133	0.7859	0.0004	0.799	0.76	0.65	0.0000	(1.06)	(1.71)
	2002	0.2524	0.0209	0.5176	0.130	0.032	(0.27)	(0.0147)	1.2205	0.0048	0.0133	0.7859	0.0097	0.790	0.75	0.62	0.0000	(1.06)	(1.68)
	2003	0.8120	0.0188	0.5501	0.162	0.330	(0.27)	(0.0168)	-0.1798	(0.0008)	0.0133	0.7859	(0.0016)	0.801	1.06	1.29	0.0000	(1.06)	(2.35)
	2004	1.9775	0.0286	0.3991	0.011	1.955	(0.27)	(0.0070)	-0.0292	(0.0001)	0.0133	0.7859	(0.0001)	0.799	0.48	0.03	0.0000	(1.06)	(1.09)
	2005	2.1303	0.0191	0.4578	0.070	2.269	(0.27)	(0.0165)	0.2895	0.0013	0.0133	0.7859	0.0026	0.797	0.29	(0.37)	0.0000	(1.06)	(0.69)
	2006	-0.4808	0.0197	0.3583	(0.030)	0.116	(0.27)	(0.0159)	-0.2307	(0.0010)	0.0133	0.7859	(0.0020)	0.801	(0.29)	(1.61)	0.0000	(1.06)	0.55
	2007	0.4568	0.0339	0.5096	0.122	0.104	(0.27)	(0.0017)	-0.6142	(0.0003)	0.0133	0.7859	(0.0006)	0.800	0.88	0.90	0.0327	0.53	(0.37)
	2008	-0.8265	0.0373	0.4183	0.030	0.342	(0.27)	0.0017	0.8440	(0.0004)	0.0133	0.7859	(0.0008)	0.800	(1.02)	(3.19)	0.0325	0.52	3.71
	2009	0.4674	0.0252	0.4615	0.073	0.109	(0.27)	(0.0104)	0.0327	0.0001	0.0133	0.7859	0.0002	0.799	0.90	0.94	0.0286	0.34	(0.61)
	2010	-0.0137	0.0374	0.2758	(0.112)	0.000	(0.27)	0.0018	-0.5646	0.0003	0.0133	0.7859	0.0006	0.799	0.42	(0.09)	0.0000	(1.06)	(0.97)
	2011	-0.1101	0.0476	0.3097	(0.078)	0.006	(0.27)	0.0120	1.1362	(0.0037)	0.0133	0.7859	(0.0074)	0.807	0.27	(0.41)	0.0461	1.18	1.59
	2012	0.0361	0.0442	0.2719	(0.116)	0.001	(0.27)	0.0086	0.3179	(0.0007)	0.0133	0.7859	(0.0015)	0.801	0.47	0.02	0.0474	1.25	1.23
	2013	1.7768	0.0422	0.2734	(0.115)	1.578	(0.27)	0.0066	0.1408	(0.0003)	0.0133	0.7859	(0.0005)	0.800	0.68	0.47	0.0248	0.15	(0.32)
	2014	0.2873	0.0415	0.2954	(0.093)	0.041	(0.27)	0.0059	-0.0714	0.0001	0.0133	0.7859	0.0002	0.799	0.74	0.60	0.0230	0.06	(0.54)
	2015	-0.1043	0.0405	0.2601	(0.128)	0.005	(0.27)	0.0049	0.2717	(0.0004)	0.0133	0.7859	(0.0007)	0.800	0.30	(0.36)	0.0374	0.76	1.12
	2016	0.1265	0.0504	0.2329	(0.155)	0.008	(0.27)	0.0148	-0.1599	0.0006	0.0133	0.7859	0.0013	0.798	0.57	0.24	0.0440	1.08	0.85
	2017	0.1113	0.0516	0.2479	(0.140)	0.006	(0.27)	0.0160	-0.1599	0.0007	0.0133	0.7859	0.0014	0.798	0.55	0.20	0.0418	0.97	0.78
	2018	0.0749	0.0666	0.2471	(0.141)	0.003	(0.27)	0.0310	-0.1599	0.0013	0.0133	0.7859	0.0027	0.797	0.50	0.07	0.0553	1.63	1.56

Table A4. Calculations of Optimal- and Excess-debt: Hong Leong Islamic Bank

	Year	Capital Gains/ (losses), (r)	Interest Rate (i)	Beta (Productivity of Capital, β)	Beta variance (αγ(0))	Half Square of capital gain variance	Correlation of interest and capital gain variables	Interest rate variance	Capital gain variance	Correlation and variances of interest and capital gain	Std. deviation of interest rate	Std. deviation of capital gain	2 x (correlation and variances of interest and capital gain)	Risk	Optimal debt ratio, $f^*(t)$	Normalized Optimal Debt ratio	Actual Debt ratio	Normalized Actual Debt ratio	Excess Debt
Hong Leong Islamic Bank	2000	0.0096	0.0258	0.5835	0.216	0.000	0.15	(0.0098)	0.3260	(0.0005)	0.0133	0.2561	(0.0010)	0.270	1.30	(0.45)	0.0057	(0.75)	(0.30)
	2001	0.1144	0.0243	0.4867	0.119	0.007	0.15	(0.0113)	0.0636	(0.0001)	0.0133	0.2561	(0.0002)	0.270	1.67	0.09	0.0000	(1.05)	(1.15)
	2002	0.6370	0.0209	0.4504	0.083	0.203	0.15	(0.0147)	1.2205	(0.0027)	0.0133	0.2561	(0.0054)	0.275	2.83	1.78	0.0075	(0.65)	(2.43)
	2003	-0.1364	0.0188	0.4679	0.100	0.009	0.15	(0.0168)	-0.1798	0.0005	0.0133	0.2561	0.0009	0.268	0.76	(1.24)	0.0007	(1.02)	0.22
	2004	0.0476	0.0286	0.3300	(0.038)	0.001	0.15	(0.0070)	-0.0292	0.0000	0.0133	0.2561	0.0001	0.269	1.43	(0.26)	0.0062	(0.73)	(0.47)
	2005	0.1459	0.0191	0.2556	(0.112)	0.011	0.15	(0.0165)	0.2895	(0.0007)	0.0133	0.2561	(0.0014)	0.271	1.78	0.26	0.0072	(0.67)	(0.92)
	2006	-0.0139	0.0197	0.3151	(0.052)	0.000	0.15	(0.0159)	-0.2307	0.0006	0.0133	0.2561	0.0011	0.268	1.25	(0.53)	0.0116	(0.43)	0.09
	2007	0.3373	0.0339	0.4206	0.053	0.057	0.15	(0.0017)	-0.6142	0.0002	0.0133	0.2561	0.0003	0.269	2.28	0.98	0.0094	(0.55)	(1.54)
	2008	-0.0595	0.0373	0.4176	0.050	0.002	0.15	0.0017	0.8440	0.0002	0.0133	0.2561	0.0004	0.269	1.00	(0.88)	0.0087	(0.59)	0.29
	2009	-0.0942	0.0252	0.5064	0.139	0.004	0.15	(0.0104)	0.0327	(0.0001)	0.0133	0.2561	(0.0001)	0.270	0.90	(1.03)	0.0092	(0.56)	0.46
	2010	0.6427	0.0374	0.3539	(0.014)	0.207	0.15	0.0018	-0.5646	(0.0002)	0.0133	0.2561	(0.0003)	0.270	2.84	1.80	0.0077	(0.64)	(2.44)
	2011	0.6704	0.0476	0.2554	(0.112)	0.225	0.15	0.0120	1.1362	0.0021	0.0133	0.2561	0.0042	0.265	2.89	1.87	0.0418	1.18	(0.70)
	2012	0.1200	0.0442	0.3313	(0.036)	0.007	0.15	0.0086	0.3179	0.0004	0.0133	0.2561	0.0008	0.269	1.63	0.03	0.0523	1.74	1.71
	2013	0.0662	0.0422	0.3155	(0.052)	0.002	0.15	0.0066	0.1408	0.0001	0.0133	0.2561	0.0003	0.269	1.45	(0.23)	0.0504	1.64	1.87
	2014	-0.0156	0.0415	0.3095	(0.058)	0.000	0.15	0.0059	-0.0714	(0.0001)	0.0133	0.2561	(0.0001)	0.270	1.15	(0.67)	0.0514	1.69	2.35
	2015	-0.1664	0.0405	0.3262	(0.041)	0.014	0.15	0.0049	0.2717	0.0002	0.0133	0.2561	0.0004	0.269	0.55	(1.55)	0.0481	1.51	3.06
	2016	0.0688	0.0504	0.2860	(0.082)	0.002	0.15	0.0148	-0.1599	(0.0004)	0.0133	0.2561	(0.0007)	0.270	1.42	(0.28)	0.0238	0.22	0.49
	2017	0.1036	0.0516	0.2923	(0.075)	0.005	0.15	0.0160	-0.1599	(0.0004)	0.0133	0.2561	(0.0008)	0.270	1.53	(0.11)	0.0160	(0.20)	(0.09)
	2018	0.2391	0.0666	0.2804	(0.087)	0.029	0.15	0.0310	-0.1599	(0.0008)	0.0133	0.2561	(0.0015)	0.271	1.89	0.40	0.0178	(0.10)	(0.51)

Table A5. Calculations of Optimal- and Excess-debt: PT Bank Mank Indocorp

	Year	Capital Gains/ (losses), (r)	Interest Rate (i)	Beta (Productivity of Capital, β)	Beta variance (σy(t))	Half Square of capital gain variance	Correlation of interest and capital gain variables	Interest rate variance	Capital gain variance	Correlation and variances of interest and capital gain	Std. deviation of interest rate	Std. deviation of capital gain	2 x (correlation and variances of interest and capital gain)	Risk	Optimal debt ratio, $f^*(t)$	Normalized Optimal Debt ratio	Actual Debt ratio	Normalized Actual Debt ratio	Excess Debt
PT Bank Mank Indocorp	2000	0.0096	0.0258	1.0100	(0.003)	0.000	(0.23)	(0.0098)	0.3260	0.0007	0.0133	0.6297	0.0015	0.641	1.55	0.13	0.1110	1.35	1.22
	2001	-0.4186	0.0243	5.3172	4.304	0.088	(0.23)	(0.0113)	0.0636	0.0002	0.0133	0.6297	0.0003	0.643	0.75	(1.40)	0.1292	1.91	3.31
	2002	0.1956	0.0209	0.8513	(0.162)	0.019	(0.23)	(0.0147)	1.2205	0.0041	0.0133	0.6297	0.0083	0.635	1.85	0.68	0.0581	(0.27)	(0.95)
	2003	1.3398	0.0188	0.9427	(0.070)	0.898	(0.23)	(0.0168)	-0.1798	(0.0007)	0.0133	0.6297	(0.0014)	0.644	2.23	1.41	0.0144	(1.60)	(3.01)
	2004	0.5260	0.0286	0.9078	(0.105)	0.138	(0.23)	(0.0070)	-0.0292	(0.0000)	0.0133	0.6297	(0.0001)	0.643	2.13	1.23	0.0126	(1.66)	(2.89)
	2005	-0.2082	0.0191	0.7268	(0.286)	0.022	(0.23)	(0.0165)	0.2895	0.0011	0.0133	0.6297	0.0022	0.641	1.19	(0.56)	0.0495	(0.53)	0.03
	2006	0.7075	0.0197	0.7680	(0.245)	0.250	(0.23)	(0.0159)	-0.2307	(0.0008)	0.0133	0.6297	(0.0017)	0.645	2.25	1.45	0.0696	0.09	(1.36)
	2007	0.1469	0.0339	0.6602	(0.353)	0.011	(0.23)	(0.0017)	-0.6142	(0.0002)	0.0133	0.6297	(0.0005)	0.643	1.73	0.47	0.1006	1.03	0.56
	2008	0.1070	0.0373	1.0879	0.075	0.006	(0.23)	0.0017	0.8440	(0.0003)	0.0133	0.6297	(0.0007)	0.644	1.67	0.35	0.0721	0.16	(0.19)
	2009	0.0655	0.0252	0.8851	(0.128)	0.002	(0.23)	(0.0104)	0.0327	0.0001	0.0133	0.6297	0.0002	0.643	1.64	0.28	0.0264	(1.24)	(1.52)
	2010	1.8075	0.0374	0.9296	(0.083)	1.633	(0.23)	0.0018	-0.5646	0.0002	0.0133	0.6297	0.0005	0.642	1.79	0.57	0.0133	(1.64)	(2.21)
	2011	-0.4712	0.0476	0.7533	(0.260)	0.111	(0.23)	0.0120	1.1362	(0.0032)	0.0133	0.6297	(0.0063)	0.649	0.59	(1.72)	0.0799	0.40	2.12
	2012	-0.0850	0.0442	0.7047	(0.308)	0.004	(0.23)	0.0086	0.3179	(0.0006)	0.0133	0.6297	(0.0013)	0.644	1.37	(0.23)	0.0882	0.65	0.89
	2013	-0.3435	0.0422	0.8080	(0.205)	0.059	(0.23)	0.0066	0.1408	(0.0002)	0.0133	0.6297	(0.0004)	0.643	0.88	(1.15)	0.0741	0.22	1.37
	2014	-0.2701	0.0415	0.6546	(0.358)	0.036	(0.23)	0.0059	-0.0714	0.0001	0.0133	0.6297	0.0002	0.643	1.03	(0.86)	0.0794	0.38	1.25
	2015	-0.2625	0.0405	0.6685	(0.344)	0.034	(0.23)	0.0049	0.2717	(0.0003)	0.0133	0.6297	(0.0006)	0.644	1.05	(0.84)	0.0590	(0.24)	0.60
	2016	1.0347	0.0504	0.5866	(0.426)	0.535	(0.23)	0.0148	-0.1599	0.0005	0.0133	0.6297	0.0011	0.642	2.28	1.50	0.0651	(0.05)	(1.56)
	2017	-0.2263	0.0516	0.5460	(0.467)	0.026	(0.23)	0.0160	-0.1599	0.0006	0.0133	0.6297	0.0012	0.642	1.11	(0.73)	0.0735	0.20	0.93
	2018	-0.1732	0.0666	0.4354	(0.577)	0.015	(0.23)	0.0310	-0.1599	0.0011	0.0133	0.6297	0.0023	0.641	1.19	(0.58)	0.0941	0.83	1.41

Table A6. Calculations of Optimal- and Excess-debt: Bank Central Asia

	Year	Capital Gains/ (losses), (r)	Interest Rate (i)	Beta (Productivity of Capital, β)	Beta variance (αγ(θ))	Half Square of capital gain variance	Correlation of interest and capital gain variables	Interest rate variance	Capital gain variance	Correlation and variances of interest and capital gain	Std. deviation of interest rate	Std. deviation of capital gain	2 x (correlation and variances of interest and capital gain)	Risk	Optimal debt ratio, $f^*(t)$	Normalized Optimal Debt ratio	Actual Debt ratio	Normalized Actual Debt ratio	Excess Debt
Bank Central Asia	2000	0.0096	0.0258	1.6925	0.812	0.000	(0.46)	(0.0098)	0.3260	0.0015	0.0133	0.3366	0.0029	0.347	2.50	(0.75)	0.0157	2.50	3.26
	2001	0.6546	0.0243	1.3646	0.484	0.214	(0.46)	(0.0113)	0.0636	0.0003	0.0133	0.3366	0.0007	0.349	3.71	1.02	0.0144	2.23	1.21
	2002	0.9932	0.0209	1.2234	0.343	0.493	(0.46)	(0.0147)	1.2205	0.0082	0.0133	0.3366	0.0165	0.333	4.10	1.58	0.0033	(0.34)	(1.92)
	2003	0.4399	0.0188	1.0072	0.126	0.097	(0.46)	(0.0168)	-0.1798	(0.0014)	0.0133	0.3366	(0.0028)	0.353	3.41	0.58	0.0016	(0.72)	(1.30)
	2004	0.6291	0.0286	0.9449	0.064	0.198	(0.46)	(0.0070)	-0.0292	(0.0001)	0.0133	0.3366	(0.0002)	0.350	3.67	0.95	0.0032	(0.35)	(1.30)
	2005	0.0796	0.0191	0.9418	0.061	0.003	(0.46)	(0.0165)	0.2895	0.0022	0.0133	0.3366	0.0044	0.346	2.72	(0.43)	0.0018	(0.69)	(0.27)
	2006	0.6682	0.0197	0.9890	0.108	0.223	(0.46)	(0.0159)	-0.2307	(0.0017)	0.0133	0.3366	(0.0034)	0.353	3.69	0.99	0.0010	(0.87)	(1.85)
	2007	0.3442	0.0339	0.8848	0.004	0.059	(0.46)	(0.0017)	-0.6142	(0.0005)	0.0133	0.3366	(0.0009)	0.351	3.22	0.31	0.0069	0.48	0.17
	2008	-0.2675	0.0373	1.0585	0.178	0.036	(0.46)	0.0017	0.8440	(0.0007)	0.0133	0.3366	(0.0014)	0.351	1.54	(2.15)	0.0103	1.26	3.41
	2009	0.7827	0.0252	0.8477	(0.033)	0.306	(0.46)	(0.0104)	0.0327	0.0002	0.0133	0.3366	0.0003	0.350	3.81	1.16	0.0065	0.39	(0.77)
	2010	0.3932	0.0374	0.7625	(0.118)	0.077	(0.46)	0.0018	-0.5646	0.0005	0.0133	0.3366	0.0010	0.349	3.32	0.45	0.0054	0.14	(0.31)
	2011	0.2276	0.0476	0.7361	(0.145)	0.026	(0.46)	0.0120	1.1362	(0.0063)	0.0133	0.3366	(0.0126)	0.363	2.84	(0.26)	0.0026	(0.51)	(0.25)
	2012	0.0834	0.0442	0.6646	(0.216)	0.003	(0.46)	0.0086	0.3179	(0.0013)	0.0133	0.3366	(0.0025)	0.352	2.60	(0.61)	0.0035	(0.30)	0.31
	2013	-0.1580	0.0422	0.7218	(0.159)	0.012	(0.46)	0.0066	0.1408	(0.0004)	0.0133	0.3366	(0.0009)	0.351	1.90	(1.61)	0.0038	(0.22)	1.39
	2014	0.3386	0.0415	0.6927	(0.188)	0.057	(0.46)	0.0059	-0.0714	0.0002	0.0133	0.3366	0.0004	0.350	3.21	0.28	0.0048	0.01	(0.27)
	2015	-0.0910	0.0405	0.6506	(0.230)	0.004	(0.46)	0.0049	0.2717	(0.0006)	0.0133	0.3366	(0.0012)	0.351	2.12	(1.30)	0.0027	(0.47)	0.83
	2016	0.1926	0.0504	0.5543	(0.326)	0.019	(0.46)	0.0148	-0.1599	0.0011	0.0133	0.3366	0.0022	0.348	2.89	(0.18)	0.0016	(0.73)	(0.56)
	2017	0.4079	0.0516	0.5127	(0.368)	0.083	(0.46)	0.0160	-0.1599	0.0012	0.0133	0.3366	0.0024	0.348	3.32	0.45	0.0008	(0.91)	(1.36)
	2018	0.1182	0.0666	0.4842	(0.397)	0.007	(0.46)	0.0310	-0.1599	0.0023	0.0133	0.3366	0.0046	0.345	2.69	(0.48)	0.0009	(0.89)	(0.42)

Table A7. Calculations of Optimal- and Excess-debt: Bank CIMB Niaga

	Year	Capital Gains/ (losses), (r)	Interest Rate (i)	Beta (Productivity of Capital, β)	Beta variance (αγ(t))	Half Square of capital gain variance	Correlation of interest and capital gain variables	Interest rate variance	Capital gain variance	Correlation and variances of interest and capital gain	Std. deviation of interest rate	Std. deviation of capital gain	2 x (correlation and variances of interest and capital gain)	Risk	Optimal debt ratio, $f^*(t)$	Normalized Optimal Debt ratio	Actual Debt ratio	Normalized Actual Debt ratio	Excess Debt
Bank CIMB Niaga	2000	0.0000	0.0258	0.0000	(0.698)	0.000	(0.28)	(0.0098)	0.3260	0.0009	0.0133	0.6074	0.0018	0.619	1.09	0.02	0.0000	(1.58)	(1.60)
	2001	0.0000	0.0243	0.0000	(0.698)	0.000	(0.28)	(0.0113)	0.0636	0.0002	0.0133	0.6074	0.0004	0.620	1.09	0.02	0.0000	(1.58)	(1.60)
	2002	0.0000	0.0209	0.9650	0.267	0.000	(0.28)	(0.0147)	1.2205	0.0050	0.0133	0.6074	0.0099	0.611	1.12	0.08	0.0380	0.37	0.29
	2003	0.0626	0.0188	0.9524	0.254	0.002	(0.28)	(0.0168)	-0.1798	(0.0008)	0.0133	0.6074	(0.0017)	0.622	1.19	0.23	0.0079	(1.17)	(1.40)
	2004	0.1977	0.0286	0.8405	0.142	0.020	(0.28)	(0.0070)	-0.0292	(0.0001)	0.0133	0.6074	(0.0001)	0.621	1.37	0.61	0.0048	(1.33)	(1.94)
	2005	0.2559	0.0191	0.8527	0.155	0.033	(0.28)	(0.0165)	0.2895	0.0013	0.0133	0.6074	0.0026	0.618	1.46	0.81	0.0256	(0.26)	(1.08)
	2006	1.5233	0.0197	1.0814	0.383	1.160	(0.28)	(0.0159)	-0.2307	(0.0010)	0.0133	0.6074	(0.0020)	0.623	1.67	1.26	0.0198	(0.56)	(1.82)
	2007	-0.0432	0.0339	0.8396	0.141	0.001	(0.28)	(0.0017)	-0.6142	(0.0003)	0.0133	0.6074	(0.0006)	0.621	1.00	(0.17)	0.0336	0.14	0.31
	2008	-0.1140	0.0373	1.0852	0.387	0.006	(0.28)	0.0017	0.8440	(0.0004)	0.0133	0.6074	(0.0008)	0.621	0.87	(0.44)	0.0380	0.37	0.82
	2009	0.7135	0.0252	0.8379	0.140	0.255	(0.28)	(0.0104)	0.0327	0.0001	0.0133	0.6074	0.0002	0.620	1.82	1.58	0.0342	0.18	(1.40)
	2010	1.8402	0.0374	0.7871	0.089	1.693	(0.28)	0.0018	-0.5646	0.0003	0.0133	0.6074	0.0006	0.620	1.30	0.48	0.0405	0.50	0.02
	2011	-0.3413	0.0476	0.7497	0.052	0.058	(0.28)	0.0120	1.1362	(0.0038)	0.0133	0.6074	(0.0076)	0.628	0.39	(1.45)	0.0379	0.37	1.82
	2012	-0.1444	0.0442	0.6455	(0.053)	0.010	(0.28)	0.0086	0.3179	(0.0008)	0.0133	0.6074	(0.0015)	0.622	0.80	(0.59)	0.0544	1.22	1.80
	2013	-0.3379	0.0422	0.6866	(0.012)	0.057	(0.28)	0.0066	0.1408	(0.0003)	0.0133	0.6074	(0.0005)	0.621	0.42	(1.39)	0.0653	1.77	3.17
	2014	-0.1114	0.0415	0.6405	(0.058)	0.006	(0.28)	0.0059	-0.0714	0.0001	0.0133	0.6074	0.0002	0.620	0.87	(0.44)	0.0649	1.75	2.19
	2015	-0.3608	0.0405	0.6620	(0.036)	0.065	(0.28)	0.0049	0.2717	(0.0004)	0.0133	0.6074	(0.0007)	0.621	0.37	(1.49)	0.0471	0.84	2.33
	2016	0.4533	0.0504	0.5870	(0.111)	0.103	(0.28)	0.0148	-0.1599	0.0007	0.0133	0.6074	0.0013	0.619	1.61	1.13	0.0263	(0.23)	(1.36)
	2017	0.5797	0.0516	0.5155	(0.183)	0.168	(0.28)	0.0160	-0.1599	0.0007	0.0133	0.6074	0.0014	0.619	1.71	1.34	0.0250	(0.29)	(1.63)
	2018	-0.3617	0.0666	0.5357	(0.162)	0.065	(0.28)	0.0310	-0.1599	0.0014	0.0133	0.6074	0.0027	0.618	0.33	(1.58)	0.0209	(0.50)	1.07

Table A8. Calculations of Optimal- and Excess-debt: Bank Mandiri

	Year	Capital Gains/ (losses), (r)	Interest Rate (i)	Beta (Productivity of Capital, β)	Beta variance (σy(t))	Half Square of capital gain variance	Correlation of interest and capital gain variables	Interest rate variance	Capital gain variance	Correlation and variances of interest and capital gain	Std. deviation of interest rate	Std. deviation of capital gain	2 x (correlation and variances of interest and capital gain)	Risk	Optimal debt ratio, $f^*(t)$	Normalized Optimal Debt ratio	Actual Debt ratio	Normalized Actual Debt ratio	Excess Debt
Bank Mandiri	2000	0.0000	0.0258	0.8687	0.137	0.000	(0.23)	(0.0098)	0.3260	0.0007	0.0133	0.5262	0.0015	0.538	1.31	(0.18)	0.1103	2.56	2.73
	2001	0.0000	0.0243	1.1538	0.423	0.000	(0.23)	(0.0113)	0.0636	0.0002	0.0133	0.5262	0.0003	0.539	1.31	(0.18)	0.1008	2.21	2.39
	2002	0.0000	0.0209	1.1206	0.389	0.000	(0.23)	(0.0147)	1.2205	0.0041	0.0133	0.5262	0.0083	0.531	1.35	(0.12)	0.0741	1.24	1.36
	2003	0.0000	0.0188	0.7271	(0.004)	0.000	(0.23)	(0.0168)	-0.1798	(0.0007)	0.0133	0.5262	(0.0014)	0.541	1.32	(0.17)	0.0550	0.55	0.72
	2004	0.7583	0.0286	0.6037	(0.128)	0.287	(0.23)	(0.0070)	-0.0292	(0.0000)	0.0133	0.5262	(0.0001)	0.540	2.17	1.42	0.0347	(0.19)	(1.61)
	2005	-0.1913	0.0191	0.7214	(0.010)	0.018	(0.23)	(0.0165)	0.2895	0.0011	0.0133	0.5262	0.0022	0.537	0.94	(0.87)	0.0347	(0.19)	0.68
	2006	0.9704	0.0197	0.7885	0.057	0.471	(0.23)	(0.0159)	-0.2307	(0.0008)	0.0133	0.5262	(0.0017)	0.541	2.24	1.54	0.0422	0.08	(1.46)
	2007	0.1623	0.0339	0.6997	(0.032)	0.013	(0.23)	(0.0017)	-0.6142	(0.0002)	0.0133	0.5262	(0.0005)	0.540	1.57	0.29	0.0122	(1.01)	(1.30)
	2008	-0.5165	0.0373	0.9185	0.187	0.133	(0.23)	0.0017	0.8440	(0.0003)	0.0133	0.5262	(0.0007)	0.540	0.08	(2.46)	0.0188	(0.77)	1.69
	2009	1.7812	0.0252	0.7547	0.023	1.586	(0.23)	(0.0104)	0.0327	0.0001	0.0133	0.5262	0.0002	0.539	1.67	0.49	0.0266	(0.49)	(0.97)
	2010	0.6247	0.0374	0.7911	0.060	0.195	(0.23)	0.0018	-0.5646	0.0002	0.0133	0.5262	0.0005	0.539	2.08	1.26	0.0151	(0.90)	(2.16)
	2011	0.0198	0.0476	0.6436	(0.088)	0.000	(0.23)	0.0120	1.1362	(0.0032)	0.0133	0.5262	(0.0063)	0.546	1.28	(0.23)	0.0122	(1.01)	(0.78)
	2012	0.1387	0.0442	0.6062	(0.125)	0.010	(0.23)	0.0086	0.3179	(0.0006)	0.0133	0.5262	(0.0013)	0.541	1.51	0.19	0.0267	(0.48)	(0.67)
	2013	-0.2327	0.0422	0.7117	(0.020)	0.027	(0.23)	0.0066	0.1408	(0.0002)	0.0133	0.5262	(0.0004)	0.540	0.79	(1.14)	0.0258	(0.51)	0.63
	2014	0.3439	0.0415	0.6352	(0.096)	0.059	(0.23)	0.0059	-0.0714	0.0001	0.0133	0.5262	0.0002	0.539	1.81	0.74	0.0262	(0.50)	(1.24)
	2015	-0.2299	0.0405	0.6381	(0.093)	0.026	(0.23)	0.0049	0.2717	(0.0003)	0.0133	0.5262	(0.0006)	0.540	0.80	(1.12)	0.0316	(0.30)	0.82
	2016	0.2806	0.0504	0.5404	(0.191)	0.039	(0.23)	0.0148	-0.1599	0.0005	0.0133	0.5262	0.0011	0.538	1.71	0.57	0.0314	(0.31)	(0.88)
	2017	0.3774	0.0516	0.5007	(0.231)	0.071	(0.23)	0.0160	-0.1599	0.0006	0.0133	0.5262	0.0012	0.538	1.83	0.79	0.0382	(0.06)	(0.85)
	2018	-0.1317	0.0666	0.4707	(0.261)	0.009	(0.23)	0.0310	-0.1599	0.0011	0.0133	0.5262	0.0023	0.537	0.98	(0.80)	0.0428	0.10	0.90

Table A9. Calculations of Optimal- and Excess-debt: Maybank Indonesia

	Year	Capital Gains/ (losses), (r)	Interest Rate (i)	Beta (Productivity of Capital, β)	Beta variance (σy(t))	Half Square of capital gain variance	Correlation of interest and capital gain variables	Interest rate variance	Capital gain variance	Correlation and variances of interest and capital gain	Std. deviation of interest rate	Std. deviation of capital gain	2 x (correlation and variances of interest and capital gain)	Risk	Optimal debt ratio, $f^*(t)$	Normalized Optimal Debt ratio	Actual Debt ratio	Normalized Actual Debt ratio	Excess Debt
Maybank Indonesia	2000	0.0096	0.0258	1.0100	(0.003)	0.000	(0.23)	(0.0098)	0.3260	0.0007	0.0133	0.6297	0.0015	0.641	1.55	0.13	0.1110	1.35	1.22
	2001	-0.4186	0.0243	5.3172	4.304	0.088	(0.23)	(0.0113)	0.0636	0.0002	0.0133	0.6297	0.0003	0.643	0.75	(1.40)	0.1292	1.91	3.31
	2002	0.1956	0.0209	0.8513	(0.162)	0.019	(0.23)	(0.0147)	1.2205	0.0041	0.0133	0.6297	0.0083	0.635	1.85	0.68	0.0581	(0.27)	(0.95)
	2003	1.3398	0.0188	0.9427	(0.070)	0.898	(0.23)	(0.0168)	-0.1798	(0.0007)	0.0133	0.6297	(0.0014)	0.644	2.23	1.41	0.0144	(1.60)	(3.01)
	2004	0.5260	0.0286	0.9078	(0.105)	0.138	(0.23)	(0.0070)	-0.0292	(0.0000)	0.0133	0.6297	(0.0001)	0.643	2.13	1.23	0.0126	(1.66)	(2.89)
	2005	-0.2082	0.0191	0.7268	(0.286)	0.022	(0.23)	(0.0165)	0.2895	0.0011	0.0133	0.6297	0.0022	0.641	1.19	(0.56)	0.0495	(0.53)	0.03
	2006	0.7075	0.0197	0.7680	(0.245)	0.250	(0.23)	(0.0159)	-0.2307	(0.0008)	0.0133	0.6297	(0.0017)	0.645	2.25	1.45	0.0696	0.09	(1.36)
	2007	0.1469	0.0339	0.6602	(0.353)	0.011	(0.23)	(0.0017)	-0.6142	(0.0002)	0.0133	0.6297	(0.0005)	0.643	1.73	0.47	0.1006	1.03	0.56
	2008	0.1070	0.0373	1.0879	0.075	0.006	(0.23)	0.0017	0.8440	(0.0003)	0.0133	0.6297	(0.0007)	0.644	1.67	0.35	0.0721	0.16	(0.19)
	2009	0.0655	0.0252	0.8851	(0.128)	0.002	(0.23)	(0.0104)	0.0327	0.0001	0.0133	0.6297	0.0002	0.643	1.64	0.28	0.0264	(1.24)	(1.52)
	2010	1.8075	0.0374	0.9296	(0.083)	1.633	(0.23)	0.0018	-0.5646	0.0002	0.0133	0.6297	0.0005	0.642	1.79	0.57	0.0133	(1.64)	(2.21)
	2011	-0.4712	0.0476	0.7533	(0.260)	0.111	(0.23)	0.0120	1.1362	(0.0032)	0.0133	0.6297	(0.0063)	0.649	0.59	(1.72)	0.0799	0.40	2.12
	2012	-0.0850	0.0442	0.7047	(0.308)	0.004	(0.23)	0.0086	0.3179	(0.0006)	0.0133	0.6297	(0.0013)	0.644	1.37	(0.23)	0.0882	0.65	0.89
	2013	-0.3435	0.0422	0.8080	(0.205)	0.059	(0.23)	0.0066	0.1408	(0.0002)	0.0133	0.6297	(0.0004)	0.643	0.88	(1.15)	0.0741	0.22	1.37
	2014	-0.2701	0.0415	0.6546	(0.358)	0.036	(0.23)	0.0059	-0.0714	0.0001	0.0133	0.6297	0.0002	0.643	1.03	(0.86)	0.0794	0.38	1.25
	2015	-0.2625	0.0405	0.6685	(0.344)	0.034	(0.23)	0.0049	0.2717	(0.0003)	0.0133	0.6297	(0.0006)	0.644	1.05	(0.84)	0.0590	(0.24)	0.60
	2016	1.0347	0.0504	0.5866	(0.426)	0.535	(0.23)	0.0148	-0.1599	0.0005	0.0133	0.6297	0.0011	0.642	2.28	1.50	0.0651	(0.05)	(1.56)
	2017	-0.2263	0.0516	0.5460	(0.467)	0.026	(0.23)	0.0160	-0.1599	0.0006	0.0133	0.6297	0.0012	0.642	1.11	(0.73)	0.0735	0.20	0.93
	2018	-0.1732	0.0666	0.4354	(0.577)	0.015	(0.23)	0.0310	-0.1599	0.0011	0.0133	0.6297	0.0023	0.641	1.19	(0.58)	0.0941	0.83	1.41

Table A10. Calculations of Optimal- and Excess-debt: Bank OCBC NISP

	Year	Capital Gains/ (losses), (r)	Interest Rate (i)	Beta (Productivity of Capital, β)	Beta variance (αγ(t))	Half Square of capital gain variance	Correlation of interest and capital gain variables	Interest rate variance	Capital gain variance	Correlation and variances of interest and capital gain	Std. deviation of interest rate	Std. deviation of capital gain	2 x (correlation and variances of interest and capital gain)	Risk	Optimal debt ratio, $f^*(t)$	Normalized Optimal Debt ratio	Actual Debt ratio	Normalized Actual Debt ratio	Excess Debt
Bank OCBC NISP	2000	0.0096	0.0258	0.8277	0.184	0.000	(0.39)	(0.0098)	0.3260	0.0013	0.0133	0.7017	0.0025	0.712	0.88	(0.08)	0.1509	3.05	3.13
	2001	-0.2543	0.0243	0.8328	0.189	0.032	(0.39)	(0.0113)	0.0636	0.0003	0.0133	0.7017	0.0006	0.714	0.47	(0.89)	0.0819	1.10	1.99
	2002	2.6166	0.0209	0.7102	0.066	3.423	(0.39)	(0.0147)	1.2205	0.0070	0.0133	0.7017	0.0141	0.701	(0.25)	(2.29)	0.0799	1.04	3.33
	2003	1.0175	0.0188	0.8278	0.184	0.518	(0.39)	(0.0168)	-0.1798	(0.0012)	0.0133	0.7017	(0.0024)	0.717	1.57	1.25	0.0853	1.20	(0.05)
	2004	0.9266	0.0286	0.6341	(0.010)	0.429	(0.39)	(0.0070)	-0.0292	(0.0001)	0.0133	0.7017	(0.0002)	0.715	1.56	1.23	0.0575	0.41	(0.82)
	2005	0.1191	0.0191	0.7660	0.122	0.007	(0.39)	(0.0165)	0.2895	0.0019	0.0133	0.7017	0.0038	0.711	1.04	0.22	0.0528	0.28	0.06
	2006	0.2077	0.0197	0.8284	0.185	0.022	(0.39)	(0.0159)	-0.2307	(0.0014)	0.0133	0.7017	(0.0029)	0.718	1.13	0.39	0.0399	(0.09)	(0.48)
	2007	0.1943	0.0339	0.6703	0.027	0.019	(0.39)	(0.0017)	-0.6142	(0.0004)	0.0133	0.7017	(0.0008)	0.716	1.10	0.33	0.0166	(0.75)	(1.08)
	2008	-0.3549	0.0373	0.8653	0.222	0.063	(0.39)	0.0017	0.8440	(0.0006)	0.0133	0.7017	(0.0012)	0.716	0.26	(1.29)	0.0175	(0.72)	0.57
	2009	0.7066	0.0252	0.6880	0.044	0.250	(0.39)	(0.0104)	0.0327	0.0001	0.0133	0.7017	0.0003	0.715	1.50	1.13	0.0145	(0.81)	(1.94)
	2010	0.7949	0.0374	0.5805	(0.063)	0.316	(0.39)	0.0018	-0.5646	0.0004	0.0133	0.7017	0.0008	0.714	1.52	1.16	0.0294	(0.39)	(1.55)
	2011	-0.2444	0.0476	0.6191	(0.025)	0.030	(0.39)	0.0120	1.1362	(0.0054)	0.0133	0.7017	(0.0108)	0.726	0.44	(0.95)	0.0295	(0.38)	0.57
	2012	0.6320	0.0442	0.5040	(0.140)	0.200	(0.39)	0.0086	0.3179	(0.0011)	0.0133	0.7017	(0.0022)	0.717	1.44	1.00	0.0111	(0.90)	(1.90)
	2013	-0.1459	0.0422	0.4666	(0.177)	0.011	(0.39)	0.0066	0.1408	(0.0004)	0.0133	0.7017	(0.0007)	0.716	0.62	(0.59)	0.0389	(0.12)	0.47
	2014	0.0826	0.0415	0.4601	(0.184)	0.003	(0.39)	0.0059	-0.0714	0.0002	0.0133	0.7017	0.0003	0.715	0.95	0.06	0.0317	(0.32)	(0.38)
	2015	-0.1590	0.0405	0.4818	(0.162)	0.013	(0.39)	0.0049	0.2717	(0.0005)	0.0133	0.7017	(0.0011)	0.716	0.60	(0.63)	0.0231	(0.56)	0.06
	2016	0.6614	0.0504	0.5110	(0.133)	0.219	(0.39)	0.0148	-0.1599	0.0009	0.0133	0.7017	0.0019	0.713	1.45	1.03	0.0237	(0.55)	(1.58)
	2017	-0.0974	0.0516	0.4656	(0.178)	0.005	(0.39)	0.0160	-0.1599	0.0010	0.0133	0.7017	0.0020	0.713	0.69	(0.46)	0.0151	(0.79)	(0.33)
	2018	-0.1410	0.0666	0.4912	(0.153)	0.010	(0.39)	0.0310	-0.1599	0.0020	0.0133	0.7017	0.0039	0.711	0.60	(0.63)	0.0184	(0.70)	(0.07)

Table A11. Calculations of Optimal- and Excess-debt: Bank Rakyat Indonesia

	Year	Capital Gains/ (losses), (r)	Interest Rate (i)	Beta (Productivity of Capital, β)	Beta variance (αγ(t))	Half Square of capital gain variance	Correlation of interest and capital gain variables	Interest rate variance	Capital gain variance	Correlation and variances of interest and capital gain	Std. deviation of interest rate	Std. deviation of capital gain	2 x (correlation and variances of interest and capital gain)	Risk	Optimal debt ratio, $f^*(t)$	Normalized Optimal Debt ratio	Actual Debt ratio	Normalized Actual Debt ratio	Excess Debt
Bank Rakyat Indonesia	2000	0.0000	0.0258	0.0000	(0.880)	0.000	(0.20)	(0.0098)	0.3260	0.0006	0.0133	0.4291	0.0013	0.441	1.94	(0.36)	0.0605	2.17	2.53
	2001	0.0000	0.0243	0.0000	(0.880)	0.000	(0.20)	(0.0113)	0.0636	0.0001	0.0133	0.4291	0.0003	0.442	1.94	(0.36)	0.0484	1.44	1.80
	2002	0.0000	0.0209	1.5957	0.716	0.000	(0.20)	(0.0147)	1.2205	0.0036	0.0133	0.4291	0.0072	0.435	1.98	(0.29)	0.0220	(0.15)	0.14
	2003	0.0000	0.0188	1.4731	0.593	0.000	(0.20)	(0.0168)	-0.1798	(0.0006)	0.0133	0.4291	(0.0012)	0.444	1.94	(0.36)	0.0244	(0.01)	0.35
	2004	1.0869	0.0286	1.1355	0.255	0.591	(0.20)	(0.0070)	-0.0292	(0.0000)	0.0133	0.4291	(0.0001)	0.442	3.05	1.33	0.0279	0.20	(1.13)
	2005	0.0082	0.0191	1.1128	0.233	0.000	(0.20)	(0.0165)	0.2895	0.0010	0.0133	0.4291	0.0019	0.440	1.98	(0.30)	0.0254	0.05	0.36
	2006	0.9014	0.0197	1.1161	0.236	0.406	(0.20)	(0.0159)	-0.2307	(0.0007)	0.0133	0.4291	(0.0015)	0.444	3.05	1.34	0.0189	(0.34)	(1.68)
	2007	0.3794	0.0339	1.1390	0.259	0.072	(0.20)	(0.0017)	-0.6142	(0.0002)	0.0133	0.4291	(0.0004)	0.443	2.61	0.66	0.0174	(0.43)	(1.09)
	2008	-0.4869	0.0373	1.3706	0.490	0.119	(0.20)	0.0017	0.8440	(0.0003)	0.0133	0.4291	(0.0006)	0.443	0.54	(2.50)	0.0071	(1.05)	1.45
	2009	0.9983	0.0252	0.9429	0.063	0.498	(0.20)	(0.0104)	0.0327	0.0001	0.0133	0.4291	0.0001	0.442	3.06	1.36	0.0126	(0.72)	(2.08)
	2010	0.4497	0.0374	1.0671	0.187	0.101	(0.20)	0.0018	-0.5646	0.0002	0.0133	0.4291	0.0004	0.442	2.70	0.79	0.0060	(1.12)	(1.91)
	2011	0.2627	0.0476	0.9143	0.034	0.034	(0.20)	0.0120	1.1362	(0.0027)	0.0133	0.4291	(0.0055)	0.448	2.36	0.29	0.0046	(1.20)	(1.49)
	2012	-0.0229	0.0442	0.7872	(0.093)	0.000	(0.20)	0.0086	0.3179	(0.0005)	0.0133	0.4291	(0.0011)	0.443	1.83	(0.52)	0.0040	(1.24)	(0.71)
	2013	-0.1741	0.0422	0.8617	(0.018)	0.015	(0.20)	0.0066	0.1408	(0.0002)	0.0133	0.4291	(0.0004)	0.443	1.46	(1.08)	0.0100	(0.88)	0.20
	2014	0.5733	0.0415	0.7272	(0.153)	0.164	(0.20)	0.0059	-0.0714	0.0001	0.0133	0.4291	0.0002	0.442	2.82	0.99	0.0147	(0.60)	(1.58)
	2015	-0.1277	0.0405	0.6979	(0.182)	0.008	(0.20)	0.0049	0.2717	(0.0003)	0.0133	0.4291	(0.0005)	0.443	1.59	(0.90)	0.0393	0.89	1.79
	2016	0.0452	0.0504	0.6082	(0.272)	0.001	(0.20)	0.0148	-0.1599	0.0005	0.0133	0.4291	0.0009	0.441	1.98	(0.30)	0.0411	1.00	1.30
	2017	0.5534	0.0516	0.5923	(0.288)	0.153	(0.20)	0.0160	-0.1599	0.0005	0.0133	0.4291	0.0010	0.441	2.79	0.93	0.0380	0.81	(0.12)
	2018	-0.0530	0.0666	0.5814	(0.299)	0.001	(0.20)	0.0310	-0.1599	0.0010	0.0133	0.4291	0.0020	0.440	1.73	(0.69)	0.0441	1.18	1.87

Table A12. Calculations of Optimal- and Excess-debt: Panin Bank

	Year	Capital Gains/ (losses), (r)	Interest Rate (i)	Beta (Productivity of Capital, β)	Beta variance (αγ(t))	Half Square of capital gain variance	Correlation of interest and capital gain variables	Interest rate variance	Capital gain variance	Correlation and variances of interest and capital gain	Std. deviation of interest rate	Std. deviation of capital gain	2 x (correlation and variances of interest and capital gain)	Risk	Optimal debt ratio, $f^*(t)$	Normalized Optimal Debt ratio	Actual Debt ratio	Normalized Actual Debt ratio	Excess Debt
Panin Bank	2000	0.0096	0.0258	0.4325	(0.076)	0.000	(0.46)	(0.0098)	0.3260	0.0015	0.0133	0.5636	0.0029	0.574	0.86	(0.19)	0.1369	3.08	3.27
	2001	0.0124	0.0243	0.6469	0.138	0.000	(0.46)	(0.0113)	0.0636	0.0003	0.0133	0.5636	0.0007	0.576	0.86	(0.18)	0.0615	(0.01)	0.17
	2002	1.8265	0.0209	0.8386	0.330	1.668	(0.46)	(0.0147)	1.2205	0.0083	0.0133	0.5636	0.0165	0.560	1.17	0.36	0.0629	0.04	(0.32)
	2003	0.6824	0.0188	0.5042	(0.005)	0.233	(0.46)	(0.0168)	-0.1798	(0.0014)	0.0133	0.5636	(0.0028)	0.580	1.62	1.17	0.0996	1.55	0.38
	2004	0.4428	0.0286	0.4813	(0.028)	0.098	(0.46)	(0.0070)	-0.0292	(0.0001)	0.0133	0.5636	(0.0002)	0.577	1.43	0.83	0.0763	0.59	(0.24)
	2005	-0.0566	0.0191	0.5202	0.011	0.002	(0.46)	(0.0165)	0.2895	0.0022	0.0133	0.5636	0.0044	0.572	0.76	(0.37)	0.0398	(0.90)	(0.53)
	2006	0.8886	0.0197	0.4192	(0.090)	0.395	(0.46)	(0.0159)	-0.2307	(0.0017)	0.0133	0.5636	(0.0034)	0.580	1.69	1.30	0.0422	(0.81)	(2.10)
	2007	0.1298	0.0339	0.3693	(0.140)	0.008	(0.46)	(0.0017)	-0.6142	(0.0005)	0.0133	0.5636	(0.0009)	0.578	1.03	0.12	0.0620	0.01	(0.11)
	2008	-0.2882	0.0373	0.6177	0.109	0.042	(0.46)	0.0017	0.8440	(0.0007)	0.0133	0.5636	(0.0014)	0.578	0.24	(1.29)	0.0628	0.04	1.33
	2009	0.8542	0.0252	0.4907	(0.018)	0.365	(0.46)	(0.0104)	0.0327	0.0002	0.0133	0.5636	0.0003	0.577	1.69	1.29	0.0571	(0.20)	(1.49)
	2010	0.5837	0.0374	0.4634	(0.046)	0.170	(0.46)	0.0018	-0.5646	0.0005	0.0133	0.5636	0.0010	0.576	1.54	1.02	0.0759	0.58	(0.45)
	2011	-0.3281	0.0476	0.4791	(0.030)	0.054	(0.46)	0.0120	1.1362	(0.0063)	0.0133	0.5636	(0.0126)	0.589	0.12	(1.51)	0.0699	0.33	1.84
	2012	-0.2336	0.0442	0.4913	(0.018)	0.027	(0.46)	0.0086	0.3179	(0.0013)	0.0133	0.5636	(0.0025)	0.579	0.35	(1.10)	0.0663	0.18	1.29
	2013	-0.1706	0.0422	0.5843	0.075	0.015	(0.46)	0.0066	0.1408	(0.0004)	0.0133	0.5636	(0.0009)	0.578	0.49	(0.86)	0.0492	(0.52)	0.34
	2014	0.7282	0.0415	0.5844	0.076	0.265	(0.46)	0.0059	-0.0714	0.0002	0.0133	0.5636	0.0004	0.576	1.61	1.16	0.0419	(0.82)	(1.98)
	2015	-0.3686	0.0405	0.4828	(0.026)	0.068	(0.46)	0.0049	0.2717	(0.0006)	0.0133	0.5636	(0.0012)	0.578	0.05	(1.63)	0.0439	(0.73)	0.90
	2016	-0.0640	0.0504	0.4535	(0.055)	0.002	(0.46)	0.0148	-0.1599	0.0011	0.0133	0.5636	0.0022	0.575	0.68	(0.50)	0.0468	(0.61)	(0.11)
	2017	0.5146	0.0516	0.4281	(0.081)	0.132	(0.46)	0.0160	-0.1599	0.0012	0.0133	0.5636	0.0024	0.574	1.46	0.89	0.0496	(0.50)	(1.39)
	2018	-0.0540	0.0666	0.3818	(0.127)	0.001	(0.46)	0.0310	-0.1599	0.0023	0.0133	0.5636	0.0046	0.572	0.68	(0.51)	0.0301	(1.30)	(0.79)

References

(Abdullah et al. 2007) Abdullah, Fikriyah, Taufiq Hassan, and Shamsher Mohamad. 2007. Investigation of Performance of Malaysian Islamic Unit Trust Funds: Comparison with Conventional Unit Trust Funds. *Managerial Finance* 33: 142–53.

(Abdul-Rahman 2010) Abdul-Rahman, Yahia. 2010. *The art of Islamic finance and Banking: Tools and Techniques for Community-Based Banking*. Hoboken: Wiley.

(Abduh, M., & Omar, M. A. 2012). Islamic banking and economic growth: the Indonesian experience. *International Journal of Islamic and middle eastern finance and management*.

(Amin and Fontaine 2013) Amin, Muslim, Zaidi Isa, and Rodrigue Fontaine. 2013. Islamic banks. *International Journal of Bank Marketing* 31: 79–97.

(Ariss 2010) Ariss, Rima Turk. 2010. Competitive conditions in Islamic and conventional banking: A global perspective. *Review of Financial Economics* 19: 101–8.

(Asutay, M., & Izhar, H. 2007). Estimating the profitability of Islamic banking: evidence from bank Muamalat Indonesia. *Review of Islamic Economics*, *11*(2), 17-29.

(Bashir 2001) Bashir, Abdel-Hameed M. 2001. Assessing the performance of Islamic banks: Some evidence from the Middle East. *Topics in Middle Eastern and North African Economies* 3: 1–11.

(Beck et al. 2010) Beck, Thorsten, Demirgüç-Kunt Asli, and Ouarda Merrouche. 2010. *Islamic vs. Conventional Banking: Business Model, Efficiency and Stability*. The World Bank. Available online: https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-5446 (accessed on 22 June 2020).

(Bitar et al. 2017) Bitar, Mohammad, Mr Sami Ben Naceur, Rym Ayadi, and Thomas Walker. 2017. *Basel Compliance and Financial Stability: Evidence from Islamic Banks*. IMF Working Papers. Washington, DC: International Monetary Fund, vol. 17.

(Brunnermeier 2009) Brunnermeier, Markus K. 2009. Deciphering the Liquidity and Credit Crunch 2007–2008. *Journal of Economic Perspectives* 23: 77–100.

(Brunnermeier et al. 2014) Brunnermeier, Markus K., and Yuliy Sannikov. 2014. A Macroeconomic Model with a Financial Sector. *American Economic Review* 104: 379–421.

(Cakranegara, P. A. 2020). Effects of Pandemic Covid 19 on Indonesia Banking. *Ilomata International Journal of Management*, *1*(4), 191-197.

(Chapra 2008) Chapra, M. Umer. 2008. The Global Financial Crisis: Can Islamic Finance Help Minimize the Severity and Frequency of Such a Crisis in the Future? Paper presented at the Forum on the Global Financial Crisis, Jeddah, Saudi Arabia, October 25.

(Eberhardt et al. 2018) Eberhardt, M. M., & Presbitero, A. (2018). *Commodity Price Movements and Banking Crises*. International Monetary Fund.

(El Said et al. 2009) El Said, Ayah, and Rachel Ziemba. 2009. *Stress-Testing Islamic Finance*. New York: Roubini Global Economics.

(Fayed 2013) Fayed, Mona Esam. 2013. Comparative performance study of conventional and Islamic banking in Egypt. *Journal of Applied Finance and Banking* 3: 1–14.

(Febianto 2012) Febianto, Irawan. 2012. Adapting risk management for profit and loss sharing financing of Islamic banks. *Modern Economy* 3: 16808.

(Gevorkyan et al. 2016) Gevorkyan, A., & Semmler, W. (2016). Oil price, overleveraging and shakeout in the shale energy sector—Game changers in the oil industry. *Economic Modelling*, *54*, 244-259.

(Gross et al. 2017) Gross, Marco, Jerome Henry, and Willi Semmler. 2017. Destabilizing Effects of Bank Overleveraging on Real Activity—An Analysis Based on a Threshold MCS-GVAR. *Macroeconomic Dynamics* 22: 1750–68.

(Hasan et al. 2010) Hasan, Maher, and Jemma Dridi. 2010. *The Effects of the Global Crisis on Islamic and Conventional Banks: A Comparative Study*. Washington, DC: International Monetary Fund.

(Hibbs 1977) Hibbs, Douglas A., Jr. 1977. Political Parties and Macroeconomic Policy. *The American Political Science Review* 71: 1467–87.

(Hidayat et al. 2012) Hidayat, Sutan Emir, and Muhammad Abduh. 2012. Does financial crisis give impacts on Bahrain Islamic banking performance? A panel regression analysis. *International Journal of Economics and Finance* 4: 79–87.

(Issa 2020). Issa, S. (2020). Life after debt: The effects of overleveraging on conventional and islamic banks. *Journal of Risk and Financial Management*, *13*(6), 137.

(Issa et al. 2022) Issa, S., & Gevorkyan, A. V. (2022). Optimal corporate leverage and speculative cycles: an empirical estimation. *Structural Change and Economic Dynamics*.

(Khan 2010) Khan, Feisal. How 'Islamic' is Islamic Banking? *Journal of Economic Behavior & Organization* 76: 805–20.

(Kusuma et al. 2016) Kusuma, Hadri, and Ariza Ayumardani. 2016. The corporate governance efficiency and Islamic bank performance: an Indonesian evidence. *Polish Journal of Management Studies* 13: 111–20.

(Marlina et al. 2021) Marlina, L., Rusydiana, A. S., Hidayat, P., & Firdaus, N. (2021). Twenty years of Islamic banking in Indonesia: A biblioshiny application. *Library Philosophy and Practice (e-journal)*, *4999*.

(Maghfuriyah et al. 2019) Maghfuriyah, A., Azam, S., & Shukri, S. (2019). Market structure and Islamic banking performance in Indonesia: An error correction model. *Management Science Letters*, *9*(9), 1407-1418.

(Mishkin 2007) Mishkin, Frederic S. 2007. *The Economics of Money, Banking, and Financial Markets*. London: Pearson education.

(Mittnik et al. 2012) Mittnik, Stefan, and Willi Semmler. 2012. Regime Dependence of the Multiplier. *Journal of Economic Behavior and Organization* 83: 502–52.

(Mittnik et al. 2013) Mittnik, Stefan, and Willi Semmler. 2013. The Real Consequences of Financial Stress. *Journal of Economic Dynamics and Control* 37: 1479–99.

(Mulyaningsih et al. 2011) Mulyaningsih, T., & Daly, A. (2011). Competitive conditions in banking industry: an empirical analysis of the consolidation, competition and concentration in the Indonesia banking industry between 2001 and 2009. *Buletin Ekonomi Moneter dan Perbankan*, *14*(2), 141-175.

(Nugroho et al. 2018) Nugroho, L., Hidayah, N., & Badawi, A. (2018). The Islamic Banking, Asset Quality: "Does Financing Segmentation Matters" (Indonesia Evidence). *Mediterranean Journal of Social Sciences*, *9*(5), 221.

(Nyambuu et al. 2015) Nyambuu, U., and L. Bernard. (2015). A Quantitative Approach to Assessing Sovereign Default Risk in Resource-Rich Emerging Economies. *International Journal of Finance & Economics* 20(3): 220–241.

(Osborn et al. 2007) Osborn, David P. J., Gus Levy, Irwin Nazareth, Irene Petersen, Amir Islam, and Michael B. King. Relative Risk of Cardiovascular and Cancer Mortality in People with Severe Mental Illness from the United Kingdom's General Practice Research Database. *Archives of General Psychiatry* 64: 242–49.

(Parashar and Venkatesh 2012) Parashar, Sat Paul, and Jyothi Venkatesh. 2012. How Did Islamic Banks Do During Global Financial Crisis. *Banks and Bank Systems* 5: 54–62.

(Pratomo and Ismail 2006) Pratomo, Wahyu Ario, and Abdul Ghafar Ismail. 2006. *Islamic Bank Performance and Capital Structure.* Munich: University Library of Munich, Germany.

(Rodoni et al. 2017) Rodoni, A., Salim, M. A., Amalia, E., & Rakhmadi, R. S. (2017). Comparing efficiency and productivity in Islamic banking: Case study Indonesia, Malaysia and Pakistan. *Al-Iqtishad: Jurnal Ilmu Ekonomi Syariah*, *9*(2), 227-242.

(Samad et al. 2004) Samad, Abdus. 2004. Performance of Interest-Free Islamic Banks Vis-À-Vis Interest-Based Conventional Banks of Bahrain. *International Journal of Economics, Management and Accounting* 12.

(Samad et al. 2006) Samad, Abdus, and M Kabir Hassan. 2006. The Performance of Malaysian Islamic Bank During 1984–1997: An Exploratory Study. *Thoughts on Economics* 10: 7–26.

(Schleer et al. 2016) Schleer, Frauke, and Willi Semmler. Banking Overleveraging and Macro Instability: A Model and VSTAR Estimations. *Jahrbücher für Nationalökonomie und Statistik* 236: 609–38.

(Shabsigh et al. 2017) Shabsigh, Ghiath, Abdullah Haron, and Mohamed Afzal Norat. 2017. In Won Song, Mariam El Hamiani Khatat, Diarmuid Murphy and Artak Harutyunyan. Ensuring Financial Stability In Countries With Islamic Banking. *International Monetary Fund* 5.

(Sole 2007) Sole, Mr Juan. 2007. *Introducing Islamic Banks into Coventional Banking Systems (Epub)*. Washington, DC: International Monetary Fund.

(Stein 2008) Stein, Jerome L. 2008. *A Tale of Two Debt Crises: A Stochastic Optimal Control Analysis*. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1092397#maincontent (accessed on 22 June 2020).

(Stein 2011) Stein, Jerome L. 2011. The Diversity of the Debt Crisis in Europe. *CESifo Working Paper No. 3348* 2011: 1–20.

(Stein 2012a) Stein, Jerome L. 2012a. Stochastic Optimal Control and the US Financial Debt Crisis. *Économie Publique/Public Economics* 26–27: 271–79.

(Stein 2012b) Stein, Jerome L. 2012b. The diversity of debt crises in Europe. In *Stochastic Optimal Control and the US Financial Debt Crisis*. Boston: Springer, pp. 133–54.

(Sufian 2014) Sufian, Fadzlan. 2014. The nexus between economic freedom and Islamic bank performance: Empirical evidence from the MENA banking sectors. *Contemporary Review of the Middle East* 1: 411–39.

(The Economist 2009) The Economist. 2009. Middle East Finance: Shine Comes Off Islamic Banks. *The Economist*, September 24.

(Turk-Ariss 2009) Turk-Ariss, Rima. 2009. Competitive Behavior in Middle East and North Africa Banking Systems. *The Quarterly Review of Economics and Finance* 49: 693–710.

(Wilson 1999) Wilson, Rodney. 1999. Challenges and Opportunities for Islamic Banking and Finance in the West: The United Kingdom Experience. *Thunderbird International Business Review* 41: 421–44.

History of the Classified Balance Sheets in Turkey by 1982

Abstract

The study is inspired by the country specific accounting history papers and aims to explore the development of the classified of balance sheets in the Ottoman and Turkey by 1982 since the Turkish Capital Market Law (1981) was first regulation that impacted the balance sheet format for public entities listed in the Istanbul Stock Exchange. The Turkish Capital Market Board developed and published the Standard Financial Statements for the Public Entities in 1984 that this first regulation of financial reporting in Turkey. For the study, secondary sources as accessible textbooks adopted by higher education institutions by 1982 in Turkey are scanned and examined. Findings of the paper might be interesting for some interested audience who can pursue comparative studies or can explore such developments and evolutions in their jurisdictions

Keywords: Balance Sheet Classification, Accounting History, Financial Reporting

JEL Code: M41, M49

Introduction

There have been currently two common approaches to present items of assets, liabilities, and owners' equity on the balance sheets for the reporting entities. The first one could be called as the Anglo-Saxon Approach (Appendix 1) that is required by FASB. The second one could be called as the Continental European Approach (Appendix 2) that is supported for some countries by the IASB. The first one lists items by the liquidity basis that is most liquid one to be first listed. On the assets side of the balance sheet, it starts with the cash account, which is most liquid asset item, then continues marketable securities, receivables, inventories, prepaid items, and other current assets. These items are grouped under the classification of current assets. Under this approach rest of the assets used to be previously classified as non-current assets, fixed assets, or capital assets. But nowadays they are classified as long-term investments, plant, property and equipment, intangible assets, and deferred assets. The first approach lists liabilities first on the right-hand side on balance sheet, then owners' equity. This comes from the liquidity basis as well. Liabilities do have maturity dates, but not owners' equity items. Also, in the legal jurisdictions, liabilities are given priorities in any case of liquidation of the entity. In this approach liabilities are classified as current and non-current or long-term depending upon their maturities.

The second approach might be interesting for people of Anglo-Saxon part of the world. Impacts and outcomes of that are seen in the Continental European Countries. It could be called as Continental European Approach because it does not follow liquidity basis. Reversely, balance sheet items are listed upon their future benefit length or productivity-based approach. They start to list goodwill and intangible assets first. Then the list goes on plant, property and equipment, long term investments, deferred assets. After those current assets are presented starting from inventories to cash account as last item of the left-hand side of balance sheet. Taking from this perspective, we could argue that this listing might be called as "futuristic or productivity-based approach" since the long future benefit providing assets are presented first. On the other hand, owners' equity items are listed first on balance sheet. In this group, capital or capital stocks are presented first. By thinking the futuristic or productivity-based approach, comparing to liabilities, owners' equity will be long lasting without having any maturity in general. Second group devotes long-term liabilities under the futuristic

approach. This group also as called as part of capital employed that is simply defined as total assets mines current liabilities (Investopedia, 2020). Last group consists of short-term liabilities.

There has been no doubt that financial information has been getting more significant day by day throughout the world due to impacts of globalization and more complex business environment. Financial reporting standard setters have been trying to protect financial information users such as investors, creditors, managers, and other stake holders particularly since the financial disasters occurred at the beginning of the millennium. As standard setters, Financial Accounting Standards Board (FASB) and International Accounting Standards Board (IASB) have worked very diligently and rigorously on the financial information including statement of financial position (balance sheet) that should be provided by the reporting entities. As direct impacts of these affords, the current formats of balance sheets are commonly seen all over the world.

In this regard, Turkey has been in a contradiction. On one hand, it is in the Southeastern Europe and has been traying to be full member of the European Union, on the other hand it has been adopting financial information format of the Anglo-Saxon Approach since 1982. Financial information on the financial statements in Turkey was first influenced by the Turkish Capital Market Law that was enacted in 1981 (Date: 28 July 1981 and Law No: 2499).

The study is inspired by certain country specific study papers (Broaker, 1896; Warner, 1899; Foulke, 1945; Iino, 1956; Walker, 1974; Edwards, 1985; Normand & Wootton, 2001; Samson, Flesher, and Previts, 2003; Walker, 2005; Yang et al.,2014). The paper attempts to find historical evidence for the history of the classified balance sheets in Turkey by 1982 in regarding changes, developments, and evolutions. After an independent war, the young Turkish Republic as the Ottoman descendant was declared on October 29, 1923. As mentioned above, a western style Turkish Capital Market Law was enacted in 1981 so that financial reporting in Turkey was impacted by the Anglo-Saxon approach.

For the study, secondary sources as accessible textbooks adopted by higher education institutions by 1982 in Turkey are scanned and examined. The paper is structured as follows. First, early examples of balance sheets existing in the Western countries are provided. Then, any evidence in the textbooks before 1928 in the Ottoman and the young Turkey is sought. Impacts of early legislation on the financial information prepared by the firms are analyzed in the first three decades of the young Turkish Republic. Thru textbooks we accessed, impacts of the influential and pioneering people in the accounting education and practice are explored and interpreted. And finally changes and developments on the classified balance sheets in Turkey by 1982 are discussed, summarized, and documented for possible further studies.

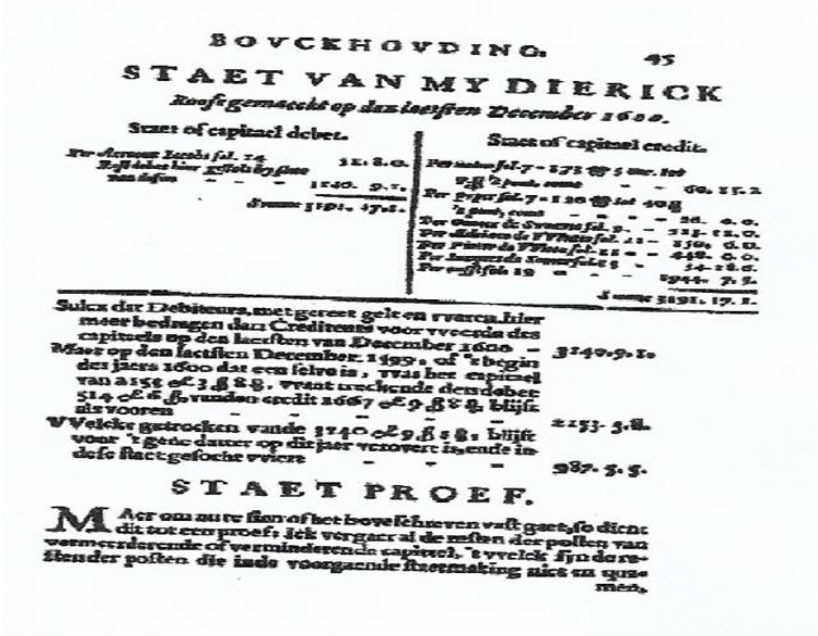
Early Examples of Balance Sheet Around the Western Countries

According to accounting historians, one of the oldest sources of statement of financial position as balance sheet is seen in the book of Indrizzo De Gli Economi by Don Angelo Pietra dated 1586 (Chatfield, 1996). The book was written for a non-profit purpose institution that it could be arguable that it first provides the entity concept that is currently included in the accounting education and practices. We could argue that items of assets of entity is listed, not grouped, or classified then, by their nature and extend on the balance sheet then.

An example of balance sheet that is like today's version of that is seen at the beginning of the Seventeenth Century. A Dutch mathematician Simon Stevin published his book of Hypomnemata Mathematica (1608) that has a formatted balance sheet (Geijsbeek, 1914, p.120; Chatfield, 1996, p.62). As seen in Exhibit 1/A and 1/B, assets are reported at the right-hand side of balance sheet and liabilities and owner's equity are reported at the left-hand side. As of today, in some countries such

as India and Egypt, balance sheets are prepared in the same format. Thus, we could argue that such balance sheet preparation is impacted by the writing style from the right to left.

Exhibit 1/A
Derrick Roose
Balance Sheet
December 31, 1600
(Source: Geijsbeek, 1914, p.120)



BOVCKHOVDING. 45

STAET VAN MY DIERICK

Staet of capitael debet. | Staet of capitael credit.

STAET PROEF.

Exhibit 1/B
Derrick Roose
Balance Sheet
December 31, 1600 (in English))
(Source: Chatfield, 1996, p.62)

TABLE 1

The Estate of Derrick Roose
made up on the last day of December 1600

Estate of Capital debit	£ s d	*Estate of Capital credit*	£ s d
(list of liabilities)	51- 8-0	(list of assets)	3191-17-1
Balance debit, to close the statement	3140- 9-1		
Total	3191-17-1	Total	3191-17-1
The remainder (Capital) at year end is			3140- 9-1
At the beginning of the year it was			2153- 3-8
Increase during the year			987- 5-5

Another best evidence might be found in India of the 18th century. As seen in Exhibit 2/A and 2/B, assets were reported on the right-hand side of balance sheet of the East India Company, and liabilities and owner's equity on the left-hand side.

Exhibit 2/A.
East İndia Company
Unclassified Balance Sheet – 1 March 1782
(Source: Baladouni, 1990, p.31)

2 (A)

Eſtimate of the General State of the Eaſt-India Company's Debts, Credits and Effects, both in England and Abroad, on the Firſt Day of March, 1782.

		Dr.
To Bonds ſtanding out bearing Intereſt	£. 1,305,500	
To ditto, depoſited in the Bank for a Loan	194,500	£. 1,497,000
To ditto, ſtanding out, not bearing Intereſt		13,154
To Bills of Exchange unpaid — from India 444,514; from China 746,401		1,190,915
To Cuſtoms on Goods ſold and unſold		1,340,562
To the Exchequer, for the laſt Payment to Government, per Agreement, for a Continuation of the Company's Right to an excluſive Trade, &c.		100,000
To Freight and Demorage		498,700
To Supra Cargoes Commiſſion on Goods ſold and unſold		70,000
To Proprietors of Private Trade		28,000
To Alms-houſes at Poplar		4,794
To Intereſt on Military and Contingent Funds more than has been applied to the Purpoſes of the ſame		93,116
To the Bank of England for Intereſt on [illegible] per Cent. per Annum		42
To Warrants paſſed the Court, unpaid		71,900
To what owing in the Departments of the Committees of Shipping		156,512
To Intereſt on Annuids		53,837
To ditto on Bonds		57,743
To Dividends on Stock unpaid, including what remains of the half Year's dividend due at Chriſtmas laſt, at £. 8 per Cent. per Annum		63,789
To what paid by the Adventurers, being £. 87 10 s. per Cent. on £. 3,200,000		2,800,000
To Balance againſt laſt Quick Stock from Bombay, the Debt due from Rugobah being deducted from the Credit of the ſame		109,794
		£. 8,141,078
To Balance		3,687,104
		£. 11,828,182

London, the 5th April, 1782.

(A) 3

Eſtimate of the General State of the Eaſt-India Company's Debts, Credits and Effects, both in England and Abroad, on the Firſt Day of March, 1782

		Cr.
By what due from Government to the Company more than the Amount of Annuities		£. 1,307,562
By Caſh, its Balance, on the firſt March, 1782		151,680
By the Company's ſeparate Fund — in Caſh £. 94,596; in Bonds 191,500		286,096
By the Amount of Goods ſold, not paid for (Diſcount deducted)		57,344
By the Value of Goods in England unſold (ditto)		3,376,619
By the prime Coſt of Goods, imported per Prime and Bellmont, not yet unladen		207,729
By what owing from the Honourable Board of Ordnance for Saltpetre delivered		270,365
By Balance of Quick Stock at Bengal, as per Abſtract, dated the 23d April, 1781. — Current Rupees 154,65,232		
Deduct what received for the Purpoſe of conſtituting a Military Fund — — 8,00,000		
Current Rupees 186,45,232		
Ditto Bellmont's Cargo arrived in England, and valued at Prime Coſt — — — 14,78,216		
Current Rupees 171,67,016		
Add Cargoes of Ships arrived at Bengal ſince the above Quick Stock, and not included therein — 8 [illegible]		
Current Rupees 180,42,958 at 2s. 3d.		8,019,831
By Balance of Quick Stock at Fort St. George, as per Abſtract, dated 9th January, 1781. — Pagodas 30,01,577		
Add Cargoes of Ships arrived at ditto ſince the above Quick Stock and not included therein — — 28,413		
Pagodas 30,29,990 at 8s.		4,011,997
By Balance of Quick Stock at Bencoolen, dated 16th October, 1780 — — Spaniſh Dollars 413,577 at 5s.		103,144
By ditto from laſt Books at St. Helena, ending 30th September, 1780. —		21,358
By ditto from ditto at China, as per Eſtimate received of Thomas Firmagh, Eſq; in November, 1781 Taels 1,097,288 at 6s. 8d.		365,763
By Cargoes of Ships ſent out in Seaſon 1780 — — —		373,758
By ditto of ditto paid for, for Exports of Seaſon 1781 — —		146,774
By Silver paid the remaining in the Treaſury in England — —		779
By Imperch paid Owners of Ships not arrived in England — —		124,367
By the Value of the Eaſt-India Houſe and Warehouſes (as eſtimated by the Company's Surveyor) — — — — —		153,616
By the Value of Ships, Sloops, and Veſſels (exclaſive of thoſe furniſhed abroad)		47,060
By what the Company paid for their dead Stock in India — —		400,000
By Expenſes defrayed for the Subſiſtence of French Priſoners in India, and incident Charges, as per General Account delivered — — —		260,687
By what remains due for Expenſes incurred in Expedition to Manilla, as per ditto		139,857
By Hoſpital Expenſes for His Majeſty's Troops at Fort St. George, Bengal and Bombay, as per ditto — — — — —		81,447
		£. 11,828,182

Exhibit 2/B
East India Company
Classified and Summarized Balance Sheet
1 March 1782
(Source: Baladouni, 1990, p. 36-37)

[Liabilities] Dr.	[Assets] Cr.
	Effective Property
Standing Debts	Standing Credits
Current Debts	Current Credits and Cash
Merchandise and Advances	Merchandise and Advances
Balance	
- Balance	*Floating Property*
	Floating Adventures
	Outward
Debts Outstanding	*Dubious Property*
	Credits Outstanding
- Balance	Doubtful Credits
	Stores
- Balance	*Dead Stock*
	Dead Stock
Balance (Total of the Balances shown above)	
[Stockholders'' Equity]	

Before 1928: The Perso-Arabic Scripts and the French Influence in the Ottoman and Turkey

The Turkish Alphabet Revolution law passed on 1 November 1928 and was effective as of 1 January 1929. By then the Ottoman Turkish alphabet that was a form of the Perso-Arabic script was mandatorily used for accounting books and financial reporting. As with Arabic, Persian, and Urdu languages, texts in the Ottoman Turkish alphabet were written right to left. It could be interesting for the western audience how to journalize the transactions in an environment of right to left writing. An example provided in the Appendix 3 could be the best evidence on the matter. In the journal entries, the format below was entirely reversed to the current one existing in the textbooks.

Account Credited	Account Debited	Credit Column	Debit Column
	Cash		xxxx
Capital Stock		xxxx	

No doubt that right to left writing impacted the format of balance sheets of the firms. In an account format of balance sheets, assets were reported on the right-hand side of balance sheets as the asset items are having debit balances as seen in Western countries previously. Consequently, liabilities and equity were reported in the left-hand side of balance sheets as they are having credit balances. There were no contra accounts of assets, liabilities, or equity items that they are listed, presented, and reported by their balances either debits or credits. An example of closing permanent or balance sheet accounts is provided in the Appendix 4/A-B. It might be evidence that during the time of right-

to-left writing periods of the Ottoman and the Republic of Turkey before the Turkish Alphabet Revolution, all balance sheets of the firms reported assets on the right-hand side of that.

It is not particular for only jurisdictions as right-to-left writing to list, present, and report assets on the right-hand side of balance sheets. As mentioned in the prior section, historically assets were presented on the right-hand side of balance sheets.

Modernized or westernized accounting education in the Ottoman was started after 1869 that the Ordinance of General Education (Maarif-i Umumiye Nizamnamesi) by the Ministry of Education. First, accounting courses (accounts and bookkeeping) were put to the curriculum of the secondary schools and high schools (Yazan & Kaya, 2017, p.174). Then, double entry bookkeeping was mandatory in 1880 for the firms in the country (Güvemli and Güvemli, 2015, p.31-32). Before 1880, double entry bookkeeping existed in the country as volunteer basis. On the other hand, before 1960 in the Ottoman there existed different bookkeeping system rather than westernized bookkeeping system. However, during the part of the 19th Century about two hundred young Turks were sent to France for westernized education. First four students named as Abdüllatif, Ahmed, Hüseyin, and Edhem were sent to France in 1830 (Erol, 2019, p.57). When they came back, they brought westernized accounting and bookkeeping techniques (Güvemli, 1997, p.4). However, before 1869 in the Ottoman accounting courses and accounting textbooks existed in the schools of minorities. Güvemli (1997, p.22) states that first accounting books having double entry bookkeeping techniques were published in the Ottoman by Armenian teachers in 1828 and 1830 (Appendix 5). Aftermath translation of French accounting textbooks to the Ottoman language continued from 1853 to 1928. Roughly forty French accounting textbooks were translated by 1928. It could be observed that the French influence had continued by the of 1930s in Turkey because Jewish-German Professors influenced not only education but also business practices till the mid of 1950s.

Some notable accounting textbooks existed in the Ottoman as following (An extended list of the textbooks complied by Oktay Güvemli (1997) is provided in the Appendix 7) :

- Fardis Efendi, Usul-i Defterî (1871) – ***Method of Bookkeeping***
- Binbasi Ziya Bey, Fenn-i Usul-i Defteri (1880) – ***Scientific Method of Bookkeeping***
- Suleyman Asaf Bey, Muhtasar Usul Defteri (1882), ***Condensed Method of Bookkeeping***
- Mehmed Mecdettin, Mebadi-i Usul-i Defterî (1884) – ***Principles of Bookkeeping Methods***
- Aram Margosyan, Muhasebe-I Ziraiye (1885) – ***Accounting for Agriculture***
- Muhyiddin, İlaveli Usul-i Defterî-i Cedid (1886) – ***Supplementary of New Bookkeeping Methods***
- Hasan Tahsin, Usul-i Defterî (1895) – ***Method of Bookkeeping***
- Hasan Tahsin, Yeni Usul-i Defterî (1895) – ***New Method of Bookkeeping***
- Mehmed Hikmet, Kavaid-i Malumat-ı Ticariye (1899) – ***Rules of Commercial Documents***
- Mehmed Hikmet, Usul-i Hesabat-ı Defteriyye (1899) – ***Method of Accounting Books***
- Ismail Hamid, Usul-i Defterî (1905) – ***Method of Bookkeeping***
- Serkiz Nihad, Defter-i Kebirli Yevmiye (1916) – ***Combined Accounting Book*** (Journal and Ledger together in a single book)
- Kirkor Kömürcüyan, Esnafın Hazır Defteri (1926) – ***Merchant's Accounting Books***

We could access Kömürcüyan's textbook of 1926 in the Ottoman language; then we observe that there do not exist financial statements as balance sheet or income statement. However, besides journal entries and ledgers, in the inventory book permanent or balance sheet accounts and their debit or credit balances, and profit/loss account exist.

Güvemli (1997) states that textbooks adopted for vocational high school education and college education by 1928 did not include financial statements as not seen in the Kömürcüyan's textbook of 1926.

The Turkish Trade Codes and Financial Reporting Before 1982

No doubt that the Turkish accounting and reporting practices were impacted by the Turkish Trade Codes. By 1982, there existed three commercial laws (1850, 1926, and 1957) to regulate companies and trading affairs in the country. The first one of 1850 called as Kanunname-i Ticaret (Yünlü, 2020, p.182) was translated from the Code de Commerce dated in 1807 by Napoleon. It was popped up among the westernization attempts. However, it was not effective well since commercial activities were regulated by the Sharia rules of the Islam (Sipahi and Küçük, 2011). It was also evidence for the French influence in the Ottoman on the accounting and reporting. Accounting practices were realized in the French language. In some cases, where ottoman language was existing, a glossary in the French language is also provided in the textbooks adopted in vocational high schools or colleges (Appendix 6) In 1926, as part of the Atatürk's Reforms, second Turkish Trade Code (Kanun-i Ticaret) passed and enacted. This law was more applicable because in 1925, the Turkish Civil Law passed and enacted that changed the commercial and non-commercial life from the Sharia rules to westernized ones.

In the first and second trade codes, there was no rules or expectations of financial reports as formatted currently. However, the accounting books were more regulated. One of the books was the inventory book that covers all permanent account and their balances. In the same book items are summarized as assets, liabilities, equity. The interpretation of the Article 75 of 1926 version of the commercial law is following:

> *The Turkish Trade Code (Kanun-i Ticaret) of 1926*
> *Article 75: Book of Accounts and Their Balances*
>
> 1- *All asset accounts and their opening balances by market values. Cash, movable and transferrable tangible assets, plant, property and equipment, accounts and notes receivable, investments on stocks, notes, and bonds.*
> 2- *All liabilities (opening balances); accounts and notes payable, and other liabilities.*
> 3- *Equity (opening balance) as the difference between assets and liabilities. In addition, all balance sheet accounts, and their ending balances must be recorded.*

Original version of Article 75 of the Turkish Trade Code (1926) by Roman letters:
YETMİŞİNCİ MADDE — Mevcudat ve muvazene defterine:
1 - Tacirin ticarete bedettiği günde malik olduğu nakit ile menkul ve gayrimenkul her nevi emvalinin kıyemi muhammenesi ve hisse senedatı ve tahvilâtıtının tarihi mezkûrda borsada cari kıymetleri ve senede merbut ve gayrimerbut bilûmum matlubatından kabili tahsil addolunan miktarı,
2 - Teahhüdat ve cihatı sairedeıı mütevellit kâffei düyunu,
3 - İşbu mevcudat ve düyunun tekabülünden hâsıl olan ve tacirin serveti asliyesini teşkil eden miktarı kayit ve işaret edildikten sonra her senei hesabiye nihayetinde bilcümle düyun ve mevcudunun mekadirini mübeyyin tanzim edeceği bilançosunu sırasile derç ve kaydetmeğe mecburdur.

As seen in the interpretation, in the first and second trade codes (1850 and 1926), recording all balance sheet items with their opening and closing balances in the inventory book was mandatory. There was no definition or requirements for financial reports as balance sheet or profit or loss statement.

The 1926 version of the Turkish Trade Code was changed in 1956, amended by the Turkish Trade Code of No: 6762 (29 June 1956). The 1956 version had a specific definition of balance sheet in the article 74 that all firms must prepare and submit to stake holders. Its interpretation is following:

> *Article 74: Balance sheet is a classified summary of the items with their balances (assets and liabilities) recorded in the inventory book.*
> *In active section of the balance sheet both tangible and intangible assets (and loss if any), in the passive side liabilities are presented.*
> *The difference between active and liabilities is the equity of the owner.*
> *The equity of the owner is presented in the passive side so that the total of the active and the total of passive will be equal. Even profit and reserves are presented separately, they are sub-items of the equity of the owner.*

Original version of Article 74 of the Turkish Trade Code (1946):
Madde 74 – *Bilanço, envanterde gösterilen kıymetlerin tasnifi ve karşılıklı olarak değerleri itibariyle tertiplenmiş hulasasıdır.*
Bilançonun aktif tablosunda; mevcutlarla alacaklar ve varsa zarar, pasif tablosunda; borçlar gösterilir.
Aktif yekünu ile borçlar arasındaki fark tacirin işletmeye tahsis ettiği ana sermayeyi teşkil eder. Ana sermaye de pasif tablosuna kaydolunur ve bu suretle aktif ve pasif tablolarının yekünları denkleşir. Yedek akçeler ve kar ayrı gösterilseler dahi, ana sermayenin cüzüleri sayılırlar.

It could be argued that this definition and requirements were adopted from the Turkish Tax Procedure Code that passed and enacted in 1949.

Impacts of the Tax Legislation on Accounting and Reporting

One could argue that accounting and reporting activities in turkey between 1950 – 1982 were more impacted by the Turkish tax legislation than the commercial ones. By 1949, there didn't exist the tax procedure law in Turkey, but different direct (individual and corporate income taxes) and indirect (property taxes, transaction taxes, etc.) taxation legislation existed. Many of those were amended by the Turkish Tax Procedure Code (Date: 7 June 1949 and Law No: 5432). It could be argued that the definition of balance sheet was first appeared in 1949 in the Tax Procedure Code (Article No: 183). As mentioned above, this was adopted by the third version of the Turkish Trade Code of 1956. The definition was so simple as listing the items by their balances of debit and credit. Items having debit balances were listed and presented on the assets as left-hand side of balance sheet. And items having credit balances were listed opposite. The interpreted version of the Article 183 is following:

> *The Turkish Tax Procedure Code (Date: 7 June 1949 and Law No: 5432)*
> *Balance Sheet*
> *Article 183: Balance sheet is a classified and summarized statement of the items listed with their balances in the accounting books.*
> *Balance sheet has two sections as active (assets) and passive (liabilities).*
> *In the active side both tangible and intangible assets (and loss if any), in the passive side liabilities are presented.*
> *The difference between active and liabilities is the equity of the owner.*
> *The equity of the owner is presented in the passive side so that the total of the active and the total of passive will be equal. Even profit and reserves are presented separately, they are sub-items of the equity of the owner.*

Original version of Article 183 of the Turkish Tax Procedure Code (1949):
Madde 183 — Bilanço, envanterde gösterilen kıymetlerin tasnifli ve karşılıklı olarak değerleri itibariyle tertiplenmiş hulâsasıdır.
Bilanço aktif ve pasif olmak üzere iki tabloyu ihtiva eder.
Aktif tablosunda mevcutlar ile alacaklar (ve varsa zarar), pasif tablosunda borçlar gösterilir.
Aktif toplamı ile borçlar arasındaki fark, müteşebbisin işletmeye mevzu varlığını (öz sermayeyi) teşkil eder.
öz sermaye pasif tablosuna kaydolunur ve bu suretle aktif ve pasif tablolarının toplamları denkleşir. İhtiyatlar ve kâr ayrı gösterilseler dahi öz sermayenin cüzüleri sayılırlar.

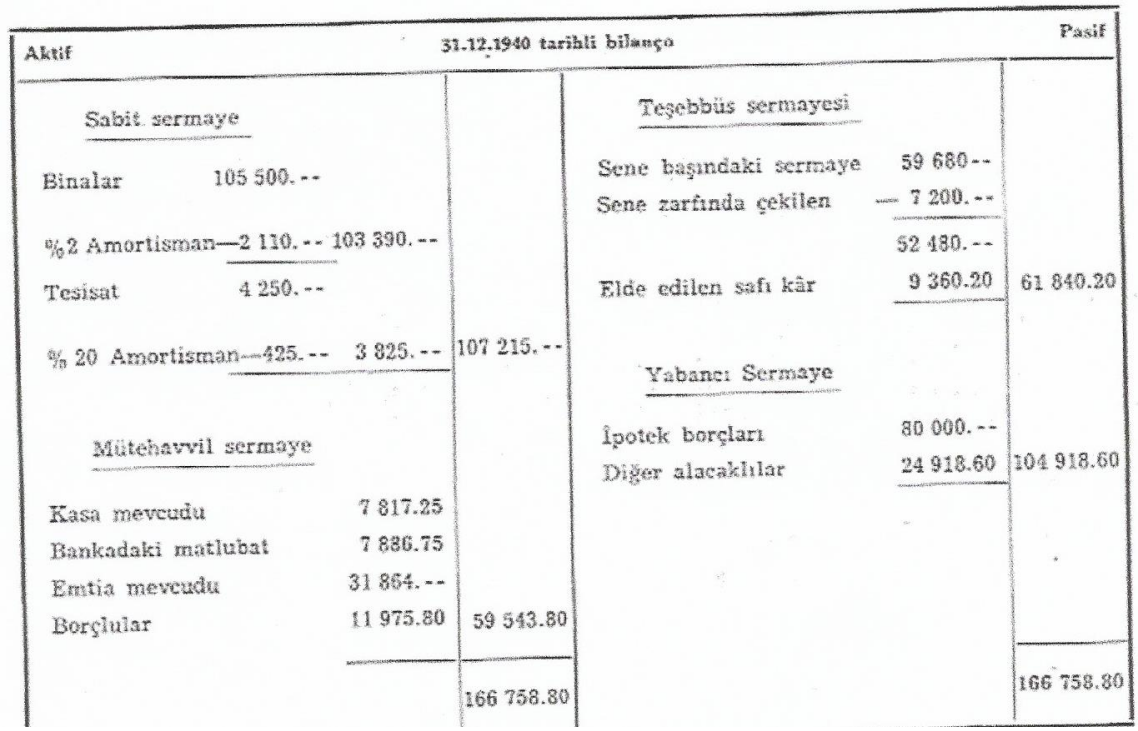
This definition of the balance sheet above remained in the second Turkish Tax Procedure Code in Article 192 (Date: 4 January 1961 and Law No: 213). Consequently, all firms that must submit their balance sheets to the Turkish fiscal authorities had to submit their balance sheets in accordance with the definition above till 1994.

Balance Sheets in the Textbook between 1928 – 1982

Kömürcüyan the Ottoman Armenian (1868 – 1958) published 44 books of business and accounting fields and continued contributions to development of accounting theory and practices in the young Turkish Republic. His first contribution in the Roman letters is the New Accounting Method (Yeni Muhasebe Usülü, 1929) that could be observed that this textbook was translated from French author Charles Rejeunne (Yaz, 2022, p.154). Unclassified balance sheet and profit or loss account are first provided in his textbook. Karataş et al. (2018, p.77) state that depreciation (amortissement in the French Language) as a concept was first introduced by Kömürcüyan in Turkey.

A classified balance sheet was first introduced by Alfred Isaac (1941) to the Turkish accounting academia and practice. Isaac (1888 – 1956) is recognized as founder of modern business education in Turkey. Isaac as German – Jewish moved to Turkey in 1937 and appointed the chair of the Department of Business Economics at the College of Economics Istanbul University till 1952. Previously, he worked for Nurnberg University, Germany (1927 – 1934) as *Ordinarius für Betriebswirtschaftslehre und Bankbetriebslehre an der Hochschule für Wirtschafts und Sozialwissenschaften Nürnberg.* Also, he resumed his career at Nurnberg University between 1952 – 1955. He published his famous book of ***Bilanzen*** (1930) in Germany, then he published first book of ***Theory of Accounting (****Muhasebe Nazariyesi****)*** (1941) that was translated to the Turkish language by Rifat Yenel. In Turkey, a classified balance sheet was seen in the Isaac's Theory of Accounting (Exhibit 3). The classification of balance sheet was oriented as the Continental European Approach mentioned in prior sections of the paper. He classified assets *(Aktif – Active)* into two major groups as ***fixed capital*** (*sabit sermaye*) meaning ***capital assets***, and ***flexible capital*** (*mütehavvil sermaye*) meaning current assets. As seen in the Exhibit 3, long-lived assets are listed and presented first. Also, contra accounts of long-lived assets first appeared on left-hand side of balance sheet in Turkey. On the other hand, it is interesting that the current assets are listed by liquidity as the Anglo-Saxon Approach. On the right-hand side *(Pasif – Passive)* of balance sheet, items of equity (*Teşebbüs sermayesi* – ***capital of the enterprise***) are listed and presented first, then liabilities (*Yabancı sermaye* – ***foreign capital***).

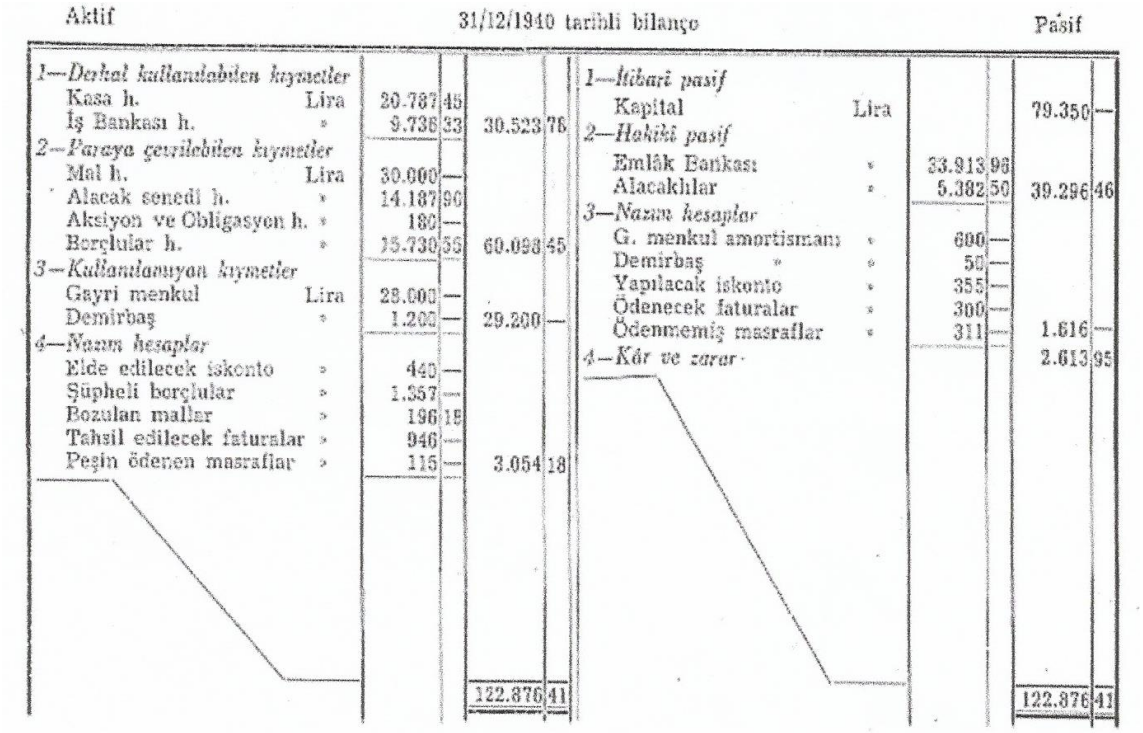
Exhibit 3: Example of Balance Sheet, Isaac (1941)



Aktif	31.12.1940 tarihli bilanço		Pasif	
Sabit sermaye			**Teşebbüs sermayesi**	
Binalar	105 500.--		Sene başındaki sermaye	59 680.--
%2 Amortisman	—2 110.--	103 390.--	Sene zarfında çekilen	— 7 200.--
				52 480.--
Tesisat	4 250.--		Elde edilen safi kâr	9 360.20 61 840.20
% 20 Amortisman	—425.--	3 825.-- 107 215.--		
			Yabancı Sermaye	
Mütehavvil sermaye			İpotek borçları	80 000.--
			Diğer alacaklılar	24 918.60 104 918.60
Kasa mevcudu	7 817.25			
Bankadaki matlubat	7 886.75			
Emtia mevcudu	31 864.--			
Borçlular	11 975.80	59 543.80		
		166 758.80		166 758.80

Cevat Yücesoy (1908 – 1969) was another pioneering scholar to introduce classified balance sheets in Turkey. Yücesoy (1941) published ***Accounting Textbook V.1*** (*Muhasebe Dersleri Cilt: 1*). It could be argued that he harmonized two approaches for balance sheet classification and presentation. Contra items of assets and liabilities were not reported, assets were listed by liquidity basis, and items on the right-hand side balance sheet were reported by the Continental European Approach (Exhibit 4). He grouped assets (*Aktif – active*) into four categories by liquidity basis as 1) ***Assets for immediate use*** (*Derhal kullanılabilen kıymetler*) as cash on hand and cash in banks, 2) ***Assets that are convertible to cash*** (Paraya çevrilebilen kıymetler) as inventories, marketable securities, notes, and accounts receivables, 3) ***Property, plant, and equipment*** (*Kullanılmayan kıymetler*) as building and equipment, 4) ***Other items*** (*Nazım hesaplar*) as interest receivable, doubtful accounts receivable, obsolete inventories, prepaid items. He grouped into four categories of items (*Pasif – passive*) reported on the right-hand side of balance sheet such as 1) ***Nominal capital*** (*İtibari sermaye*) as capital or capital stocks, 2) ***Real capital*** (Hakiki sermaye) as bank notes payable and accounts payable, 3) ***Other items*** (*Nazım hesaplar*) as accumulated depreciation, interest payable, accrued expenses, and other payables, 4) ***Profit*** (*Kar ve zarar*).

Exhibit 4: **Example of Balance Sheet, Yücesoy (1941)**



Aktif — 31/12/1940 tarihli bilanço — Pasif

1—Derhal kullanılabilen kıymetler				*1—İtibarî pasif*			
Kasa h.	Lira	20.787 45		Kapital	Lira		79.350 —
İş Bankası h.	»	9.736 33	30.523 78	*2—Hakikî pasif*			
2—Paraya çevrilebilen kıymetler				Emlâk Bankası	»	33.913 96	
Mal h.	Lira	30.000 —		Alacaklılar	»	5.382 50	39.296 46
Alacak senedi h.	»	14.187 90		*3—Nazım hesaplar*			
Aksiyon ve Obligasyon h.	»	180 —		G. menkul amortismanı	»	600 —	
Borçlular h.	»	15.730 55	60.098 45	Demirbaş	»	50 —	
3—Kullanılmayan kıymetler				Yapılacak iskonto	»	355 —	
Gayri menkul	Lira	28.000 —		Ödenecek faturalar	»	300 —	
Demirbaş	»	1.200 —	29.200 —	Ödenmemiş masraflar	»	311 —	1.616 —
4—Nazım hesaplar				*4—Kâr ve zarar*			2.613 95
Elde edilecek iskonto	»	440 —					
Şüpheli borçlular	»	1.357 —					
Bozulan mallar	»	196 19					
Tahsil edilecek faturalar	»	946 —					
Peşin ödenen masraflar	»	115 —	3.054 18				
			122.876 41				122.876 41

Besides the books mentioned above, we could access some other accounting textbooks having classified balance sheets by 1982. We summarized those in Table 1 that authors, name of the books, their publication years, approaches, or bases adopted are included. It could be argued that each book has its own concepts and classifications. It means there did not exist any uniformity that each author was impacted from different sources of different countries such as France, Germany, the United Kingdom, the United States, etc.

TABLE 1: Selected Accounting Textbooks in Turkey (1950 – 1981)

Source: Pekdemir et al., 2021, p.21-22

Author	Name of the Book	Year published	Approach Followed	Groups of Assets
Cevat Yücesoy	Envanter ve Bilanço	1957 and 1970	Continental European	**Model 1** Sabit Kıymetler Dönen Kıymetler **Model 2** Sabit Kıymetler Emre Hazır Kıymetler Paraya Çevrilebilen Kıymetler Pasifi Tanzim Edici Hesaplar Kar ve Zarar
Nevzat Alptürk	Krediyi Tayin Eden Unsurların Tahlili Rasyolar	1957	Continental European	Sabit Kıymetler Bir İşe Bağlanmış Kıymetler İşletme Kıymetleri Tahakkuk Ettirilen Kıymetler Mevcut Kıymetler
Cevat Yücesoy	Ticaret Şirketleri ve Muhasebesi	1958	Continental European	Sabit Kıymetler Emre Hazır Kıymetler Paraya Çevrilebilen Kıymetler
Sedat Ünalan	Muhasebe	1960	None But, Tax Code	Mevcutlar Alacaklar
Türkiye Bankalar Birliği Yayın No: 20	Kredi Talebedenlerden Alınacak Hesap Vaziyeti ve İnceleme Tabloları	1963	Anglo-Saxon	Mütedavil Kıymetler (Cari Aktifler) Bağlı Kıymetler Sabit Kıymetler
Cevat Yücesoy	Bilanço Tahlilleri	1968	Continental European	Sabit Kıymetler Dönen Kıymetler
Yüksel Koç Yalkın	İşletmelerde Mali Analiz Teknikleri	1970	Anglo-Saxon	Döner Değerler Uzun Vadeli Aktifler
Alp Gürus	Ticaret Bankalarımızda Finansal Analiz ve Rasyo Metodunun Uygulanması	1972	Anglo-Saxon	Döner Varlıklar Dönen Varlıklar Bağlı Varlıklar Değişmez Varlıklar
Mustafa A. Aysan	Finansal Tablolar Analizi ve Diğer Muhasebe Konuları	1976	Anglo-Saxon	Mütedavil Kıymetler (Cari Aktifler) Bağlı Kıymetler Sabit Kıymetler
Atilla Gönenli ve Cevat Sarıkamış	İşletme Finansı ve Okuma Parçaları	1976	Anglo-Saxon	Cari Aktifler Sabit Aktifler
Yapı ve Kredi Bankası	Mali Tahlil El Kitabı	1981	Anglo-Saxon	Mütedavil Kıymetler (Cari Aktifler) Bağlı Kıymetler Sabit Kıymetler Zarar

TABLE 2: Selected Accounting Textbooks (Translated from English) in Turkey (1962 – 1980)
Source: Pekdemir et al., 2021, p.21-22

Author	Name of the Book	Year published	Approach Followed	Groups of Assets
Paton, W.A. and Dixon, R. I.	Essentials of Accounting (Muhasebenin Temelleri) Çev.: Mustafa A. Aysan	1962	Anglo-Saxon	Cari Aktif Cari Olmayan Aktif
Kennedy, R.D. and McMullen, S.	Statements of Financial Position (Finansal Durum Tabloları) Çev.: Atilla Gönenli	1962	Anglo-Saxon	Cari Aktiflar Uzun Vadeli Yatırımlar Sabit Aktifler Maddi Olmayan Aktifler Gelecek Devrelere Ait Harcamalar Diğer Aktifler
Moore, C.L. and Jaedicke, R.K.	Managerial Accounting (Yönetim Muhasebesi) Çev.: Alpaslan Peker	1980	Anglo-Saxon	Dönen Varlıklar

As seen in the Table 1 and 2, in the second quarter of the 20th century the Continental European Approach was mainly adopted for the classified balance sheets since German influence exited by the mid of 1950s in Turkey. Then, the Anglo-Saxon Approach as liquidity basis was adopted later because after World War II Turkey became part of Marshall Plan and the OECD Establishment (Üstün, 1997, p.31). That means American textbooks were adopted in the institutions where the English language was adopted for the instruction. In addition, several accounting and business faculty members were sent to the States (Aysan et. Al., 2020, 9.7) to improve their proficiency in English, some others were sent for further education opportunities to attend graduate programs. On the other hand, '*assets*' were translated into different Turkish words as '*değerler*,' *'kıymetler*,' 'akt'fler,' and '*varlıklar*' in different time by different textbook authors.

It could be argued that development or evolution of the classified balance sheets in Turkey are based on the translation that different authors made different translation. Such examples were seen many years ago in western countries. Walker (1974) stated that classified balance sheets were almost uniformed in the mid of second quarter of the 20th Century in the United States and the United Kingdom. This development was realized in Turkey roughly fifty years later that first regulations on the matter were seen just after the Turkish Capital Market Law of 1981.

Summary and Conclusion

The study is inspired by the country specific accounting history papers and aims to explore the development of the classified of balance sheets in the Ottoman and Turkey by 1982 since the Turkish Capital Market Law (1981) was first regulation that impacted the balance sheet format for public entities listed in the Istanbul Stock Exchange. The Turkish Capital Market Board developed and published the Standard Financial Statements for the Public Entities in 1984. After that, accounting textbooks published in Turkey started to adopt the regulation. For the study, from earliest examples of the textbooks to 1982 are scanned and the evidence is sought. Thus, secondary sources from the late Ottoman period and first half-century of Turkey are covered. Findings might be interesting for some interested audience.

- First accounting textbooks are published by Ottoman-Armenians in 1828 and 1830.
- First students were sent to France to pursue their higher education in 1830.
- First version of the Trade Code in the Ottoman Empire was translated from the Code de Commerce of Napoleon of 1807 to the Ottoman Language and enacted in 1850.
- Ordinance of the General Education was first prepared and enacted in 1869.
- Double-entry bookkeeping was first required in 1880.
- Gregorian Calendar was first used in 1920ö then continued.
- The Republic of Turkey was declared, and the Ottoman Empire ended in 1923.
- Second version of the Turkish Trade Code enacted in 1926.
- The Turkish Alphabet changed from the Perso-Arabic script to Roman in 1928.
- French influence existed on the accounting education in the Ottoman and in Turkey by mid of 1930s.
- Jewish-German professor were invited in the mid of 1930s, then they moved to Turkey and notably impacted the Turkish higher education and society.
- Between mid of 1930s and the beginning of 1960s, German influence existed on not only accounting education but accounting practices in Turkey.
- Third version of the Turkish Trade Code enacted in 1946.
- First Turkish Tax Procedure Code enacted in 1949.
- Second version of Turkish Tax Procedure Code enacted in 1961.

From the beginning of 1960s, the Anglo-Saxon approaches started to influence on accounting, education, accounting practices, and accounting profession. It could be argued that in somewhere Americanized financial reporting is mainly adopted in Turkey even though Turkey has endorsed and adopted International Financial reporting Standards (IFRSs) developed and published by the International Accounting Standards Board.

The paper argues that classified balance sheets in Turkey by 1982 are evolved by the translation of French, German, and American sources. It might be seen in the classified balance sheets existed in the accounting textbooks published by different authors having different backgrounds.

The study has some limitations of coverage that only accessible textbooks and other secondary sources are scanned. There could be more in both the late Ottoman period and the first half century of Turkey. The authors have disabilities to read or understand the Perso-Arabic script. On the other hand, not all sources existing in the period that the study aims to cover are digitalized.

This is a country specific study paper that can contribute to accounting history for the interested audience who can pursue comparative studies or can explore such developments and evolutions in their jurisdictions.

REFERENCES

Alptürk, N. (1957). Krediyi tayin eden unsurlar bilançoların tahlili rasyolar. Ankara.

Aysan, M. A. (1976). Finansal tablolar analizi ve diğer muhasebe konuları. İstanbul: Sermet Matbası.

Aysan, M. A., Yücel, G., & Adiloğlu, B. (2020). Türkiye'de Lisansüstü İşletmecilik Eğitiminde Bir Kuyruklu Yıldız: "İşletmecilik İhtisas Programı". Muhasebe ve Finans Tarihi Araştırmaları Dergisi(19), 3-24.

Baladouni, V. (1990). An early attempt at balance sheet classification and financial reporting, The Accounting Historians Journal, Vol. 17, No. 1, (June), 27-45.

Broaker, F. (1896) (reprint 2012). The american accountants' manual, - examiation questions volume 1. London, UK: Forgotten Books.

Chatfield, M. & Vangermeersch, M. (1996). The history of accounting – an iInternational encyclopedia. New York, NY: Garland Publıshıng, Inc. Current. (2020, September 20).

Edey, H. C., & Panitpakdi, P. (1956). British company accounting and the law 1840 -1900. In A. C. Littleton & B. S. Yamey (Eds.), Studies in the history of accounting (pp. 356-379). Homewwod, IL, Richard D. Irwin Inc. (reprint in 2013, New Delhi, Isha Books, p.356-379.

Erol, S. (2019), Istanbul'dan Paris'e gönderilen ilk Oamanlı talebelerinden Edhem Efendi'nin egitim hayatı, İçtimaiyat Sosyal Bilimler Dergisi, V.3 (1), p.53-69.

Foulke, R. A. (1945). Practical financial statement analysis. New York, NY: McGraw-Hill Book Company, Inc. Geijsbeek, J. B. (1914). Ancient Double Entry Bookeeping. Denver, CO.

Gencoğlu, Ü. G. (2009), Türkiye'de Harf Devriminin Muhasebe Eğitimi ve Muhasebe Uygulamalarına Etkisi, Akademik Araştırmalar ve Çalışmalar Dergisi (AKAD) Cilt: 1 - Sayı: 1, p.30-37.

Gönenli, A. and Sarıkamış, C. (1976). İşletme finansı okuma parçaları ve vak'alar. İstanbul: Sermet Matbaası.

Gürus, A. (1972). Ticaret bankalarımızda finansal analiz ve rasyo metodunun uygulanması. Ankara: Başnur Matbaası.

Güvemli, O. (1997), 19.Yüzyılda Türkiyede Muhasebe Öğretim Kitapları, Muhasebe Öğretim Üyeleri Bilim ve Dayanışma Vakfı (MÖDAV) Bülten Ek Yayını n.:3, October, p.1-51

Güvemli, O. and Güvemli, B. (2015). "Türk Ticaret Kanunlarının Türk Muhasebe Düşüncesinin Gelişmesindeki Etkileri", Muhasebe ve Finans Tarihi Araştırmaları Dergisi, n.8, 26-50.

Iino, T. (1956), Accounting classification of assets. The Annals of the Hitotsubashi Academy, Vol. 6, No. 2, (April), 66-78.

Isaac A. (1941). Muhasebe nazariyesi (R. Yenel, Çev.). İstanbul: Güven Basımevi.

Karataş, Ş., Adiloğlu, B. and Yucel, G (2018), Muhasebe eğitimine verilen doksan yıllık ömür: Kirkor Kömürcüyan, Muhasebe Enstitüsü Dergisi, Cilt: 16, Sayı: 19, p.71-80.

Kennedy, R. D., & Mc Mullen, S. Y. (1962). Finansal durum tabloları (A. Gönenli, Çev.). İstanbul: Sermet Matbaası.

Koç Yalkın, Y. (1970). İşletmelerde mali analiz teknikleri (2. bs.). Ankara: Sevinç Matbaası.

Kömürciyan, K. (1926). Esnâfın hâzır defteri. İstanbul: Sebât Matba'ası.

Moore, C. L., & Jaedicke, R. K. (1980). Yönetim muhasebesi (A. Peker, Çev.). İstanbul: Fatih Yayınevi Matbaası.

Normand, C., & Wootton, C. W. (2001). The recognition and valuation of current assets on the balance sheet in the United States, 1865- 1940. The Accounting Historians Journal, Vol. 28, No. 2, (December), 63-108.

Paton, W. A., & Dixon, R. L. (1962). Muhasebenin temelleri (M. A. Aysan, Çev.). İstanbul: Sermet Matbaası.

Pietra, D. A. (1586). Indrizzo de gli economi (J. B. Geijsbeek, Trans.). (1914). Ancient Double Entry Bookeeping (pp. 87-107). Denver, CO. Quick. (2020, September 20).

Sipahi, B, and Küçük, I. (2011), Türk Ticaret Kanunları ve Muhasebenin Gelişimine 160 Yıllık Etkileri, MUFTAV, Oktay Güvemli Muhasebe ve Tarih Vakfı Dergisi, Sayı 1, Temmuz 2011, p.181-193

Stevin, S. (1608). Hypomnemata mathematica (J. B. Geijsbeek, Trans.). 1914, Ancient Double Entry Bookeeping, (pp. 1114-136). Denver, CO. Thornton, J. (1895) (reprint Forgatten Books 2018). A manual of bookkeeping for the use of students. London: MacMillan and Co. Limited.

Türkiye Bankalar Birliği. (1963). Kredi talebedenlerden alınacak hesap vaziyeti ve inceleme tabloları. Bankaların açacakları krediler için talep sahiplerinden alacakları hesap vaziyeti ve tahlil esasları semineri kitabı içerisinde (s. 8-17). İstanbul: Türkiye Bankalar Birliği Yayınları No: 20.

Ünalan, S. (1960). Muhasebe. Ankara: Doğuş Ltd. Şirketi Matbaası.

Üstün, S. (1997), Turkey and Marshall Plan: Strive for Aid. M.A. Thesis, presented to the Department of History, University College London, 1997. P. 31-52.

Walker, R. G. (1974). Asset classification and asset valuation, Accounting Business Research, (Autumn), 286-296.

Warner, R. (1899). '"Fixed" and "circulating" assets', correspondance, The Accountant, 4 March. (Quoted: Walker, 1974).

Yang, S., Yang, R., & Lin, Z. J. (2014). Historical evolution of the balance sheet in the people's republic of China. The Accounting Historians Journal, (June), Vol. 41, No. 1, 1-34.

Yapı ve Kredi Bankası (1981). Mali tahlil el kitabı.

Yaz, D. A. (2022), Uygarlığa Yön Veren Türk Muhasebe Filozofları, Timaş Yayınları, No 5436.

Yazan, Ö. And Kaya, M. (2017), Ali Suavi ve Muhasebe Öğretimi: Fenn-i Tanzim-i Defter (1869), Uluslararası Sosyal Araştırmalar Dergisi V.10, N.51, 1174-1180.

Yücesoy, C. (1941). Muhasebe dersleri cilt: 1. İstanbul: Kenan Basımevi ve Klişe Fabrikası.

Yücesoy, C. (1953). Muhasebe dersleri cilt: 1 (3. bs.). İstanbul: Ahmet Halit Yaşaroğlu Kitapçılık ve Kağıtçılık T. L. Ş.

Yücesoy, C. (1957). Envanter ve bilanço. İstanbul: Işıl Matbaası.

Yücesoy, C. (1958). Ticaret şirketleri ve muhasebesi. İstanbul: Işıl Matbaası.

Yücesoy, C. (1968). Bilanço tahlilleri (4. bs.). İstanbul: Çağlayan Kitabevi.

Yücesoy, C. (1970). Envanter ve bilanço (5.bs.). İstanbul: Çağlayan Kitabevi.

Yünlü, M. (2020), Türkiyede'ki Muhasebe Uygulamalarının Tarihsel Gelişimi ve Günümüzdeki Durumu, Muhasebe ve Finans Tarihi Araştırmaları Dergisi, January (18), 180-192.

Appendix 1

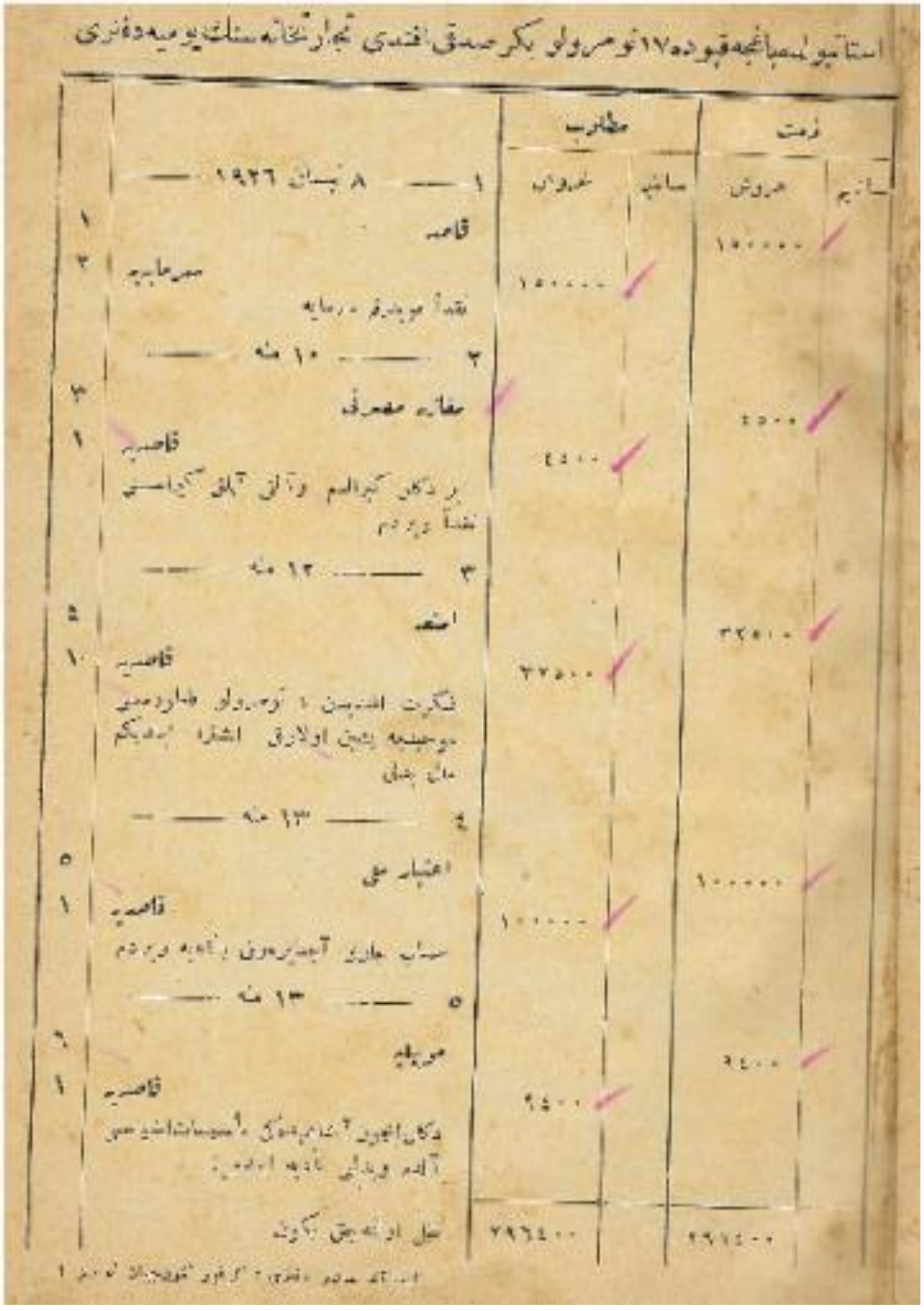
FORD MOTOR COMPANY AND SUBSIDIARIES
CONSOLIDATED BALANCE SHEETS
(in millions)

	December 31, 2019	December 31, 2020
ASSETS		
Cash and cash equivalents (Note 9)	$ 17,504	$ 25,243
Marketable securities (Note 9)	17,147	24,718
Ford Credit finance receivables, net of allowance for credit losses of $162 and $394 (Note 10)	53,651	42,401
Trade and other receivables, less allowances of $63 and $84	9,237	9,993
Inventories (Note 11)	10,786	10,808
Assets held for sale (Note 2, Note 10, and Note 22)	2,383	47
Other assets	3,339	3,534
Total current assets	114,047	116,744
Ford Credit finance receivables, net of allowance for credit losses of $351 and $911 (Note 10)	53,703	55,277
Net investment in operating leases (Note 12)	29,230	27,951
Net property (Note 13)	36,469	37,083
Equity in net assets of affiliated companies (Note 14)	2,519	4,901
Deferred income taxes (Note 7)	11,863	12,423
Other assets	10,706	12,882
Total assets	$ 258,537	$ 267,261
LIABILITIES		
Payables	$ 20,673	$ 22,204
Other liabilities and deferred revenue (Note 16 and Note 25)	22,987	23,645
Automotive debt payable within one year (Note 19)	1,445	1,194
Ford Credit debt payable within one year (Note 19)	52,371	49,969
Other debt payable within one year (Note 19)	130	180
Liabilities held for sale (Note 22)	526	—
Total current liabilities	98,132	97,192
Other liabilities and deferred revenue (Note 16 and Note 25)	25,324	28,379
Automotive long-term debt (Note 19)	13,233	22,342
Ford Credit long-term debt (Note 19)	87,658	87,708
Other long-term debt (Note 19)	470	291
Deferred income taxes (Note 7)	490	538
Total liabilities	225,307	236,450
EQUITY		
Common Stock, par value $0.01 per share (4,025 million shares issued of 6 billion authorized)	40	40
Class B Stock, par value $0.01 per share (71 million shares issued of 530 million authorized)	1	1
Capital in excess of par value of stock	22,165	22,290
Retained earnings	20,320	18,243
Accumulated other comprehensive income/(loss) (Note 23)	(7,728)	(8,294)
Treasury stock	(1,613)	(1,590)
Total equity attributable to Ford Motor Company	33,185	30,690
Equity attributable to noncontrolling interests	45	121
Total equity	33,230	30,811
Total liabilities and equity	$ 258,537	$ 267,261

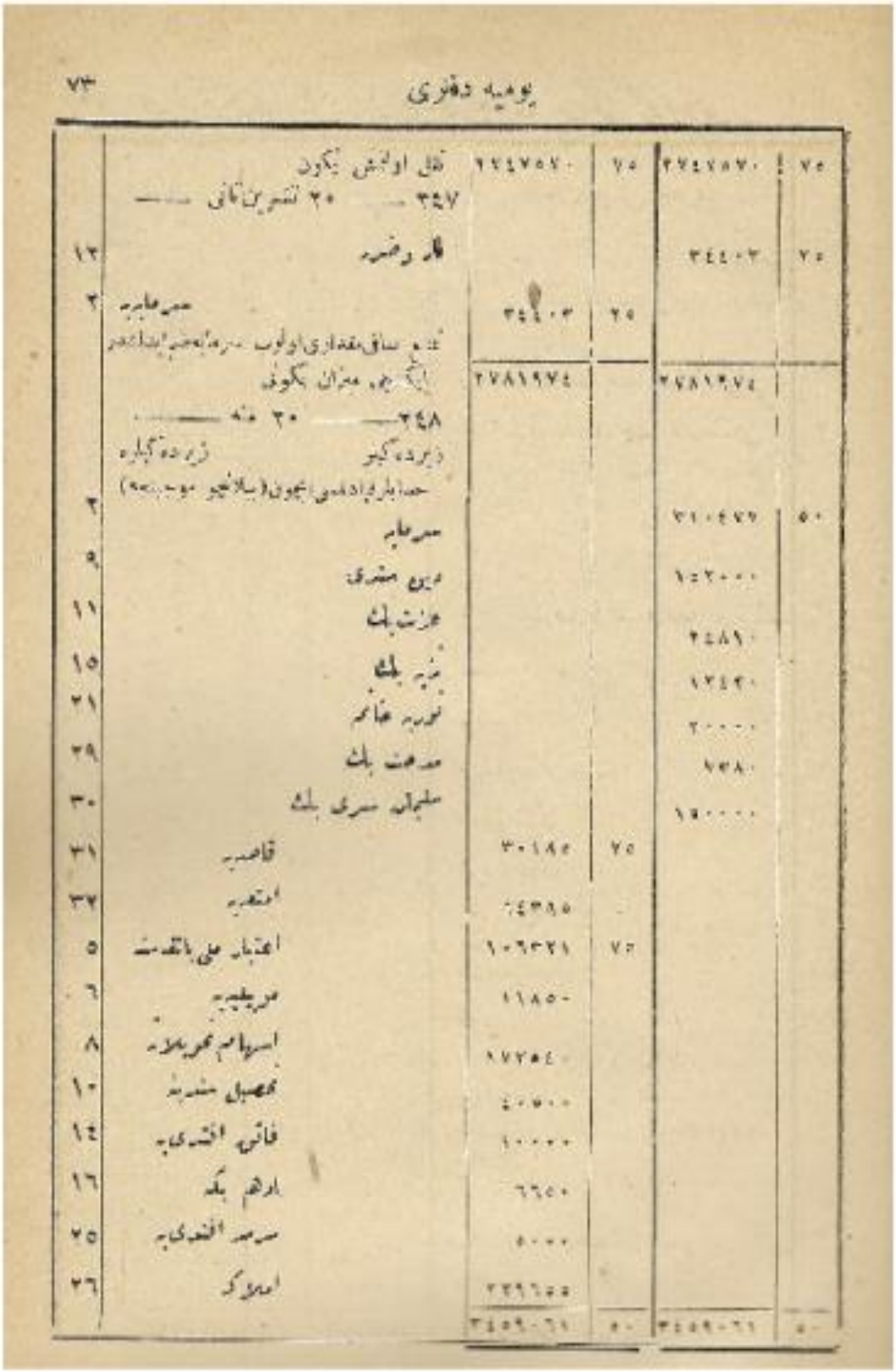
Appendix 2: Mercedes – Benz Group
Consolidated Statement of Financial Position

In millions of euros	Note	At 31 December 2021	2020
Assets			
Intangible assets	11	**15,005**	16,399
Property, plant and equipment	12	**27,859**	35,246
Equipment on operating leases	13	**44,471**	47,552
Equity-method investments	14	**13,588**	5,189
Receivables from financial services	15	**46,955**	53,709
Marketable debt securities and similar investments	16	**873**	1,041
Other financial assets	17	**3,181**	4,167
Deferred tax assets	10	**3,434**	6,259
Other assets	18	**1,536**	911
Total non-current assets		**156,902**	170,473
Inventories	19	**21,466**	26,444
Trade receivables	20	**7,673**	10,649
Receivables from financial services	15	**33,670**	42,476
Cash and cash equivalents		**23,120**	23,048
Marketable debt securities and similar investments	16	**6,706**	5,356
Other financial assets	17	**3,079**	2,757
Other assets	18	**4,073**	4,534
Assets held for sale	3	**3,142**	-
Total current assets		**102,929**	115,264
Total assets			
Equity and liabilities			
Share capital		**3,070**	3,070
Capital reserves		**11,723**	11,551
Retained earnings		**56,190**	47,111
Other reserves		**968**	-1,041
Equity attributable to shareholders of Mercedes-Benz Group AG		**71,951**	60,691
Non-controlling interests		**1,216**	1,557
Total equity	21	**73,167**	62,248
Provisions for pensions and similar obligations	23	**5,359**	12,070
Provisions for other risks	24	**7,909**	11,116
Financing liabilities	25	**73,543**	86,539
Other financial liabilities	26	**1,808**	1,971
Deferred tax liabilities	10	**4,488**	3,649
Deferred income	27	**1,175**	1,567
Contract and refund liabilities	28	**3,980**	5,787
Other liabilities	29	**727**	981
Total non-current liabilities		**98,989**	123,680
Trade payables		**10,655**	12,378
Provisions for other risks	24	**8,053**	9,334
Financing liabilities	25	**52,300**	59,303
Other financial liabilities	26	**5,997**	6,627
Deferred income	27	**1,486**	1,594
Contract and refund liabilities	28	**5,929**	7,169
Other liabilities	29	**3,086**	3,404
Liabilities held for sale	3	**169**	-
Total current liabilities		**87,675**	99,809
Total equity and liabilities		**259,831**	285,737

Appendix 3
First Page of Merchant's Accounting Textbook of Krikor Kömürcüyan (1926)



Appendix 4/A
First Page (37) of Merchant's Accounting Textbook of Krikor Kömürcüyan (1926) in Arabic Letters

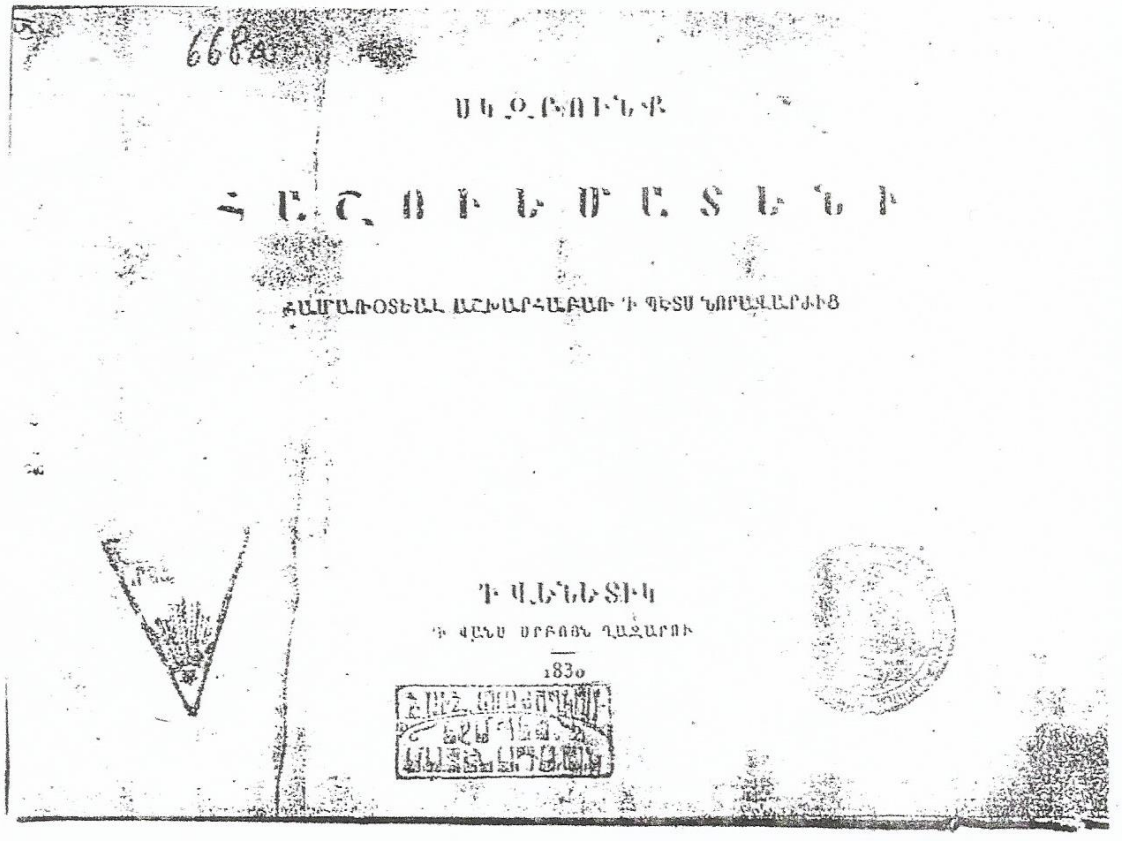


Appendix 4/B
First Page (37) of Merchant's Accounting Textbook of Krikor Kömürcüyan (1926) in Roman Letters

37	Yevmiyye Defteri					
			Matlûp		Zimmet	
			Gurûş	Santim	Gurûş	Santim
	nakl olunmuş yekûn		2747570	75	2747570	75
20 Teşrîn-i Sânî - 347						
12		Kâr ve Zarar			34403	25
2	Sermâyeye		34403	25		
Temeddî sâfî miktârı alup sermâyeye zamm olunmuştur						
	İkinci Mizân-ı yekûnı		2781974	--	2781974	--
Minhâ - 348						
	Zîrdekilere	Zîrdekiler				
Hesâblar kapadılmazdan için (Bilanço mûcebince)						
2		Sermâye			310477	50
9		Deyn senedi			152000	—
11		İzzet Beğ			24810	—
15		Nemlı Beğ			12420	—
21		Nuriye Hânım			20000	—
29		Mürsad Beğ			7380	—
30		Süleymân Sırrı Beğ			150000	—
31	Kasaya		30185	75		
32	Emtî'aya		64385	—		
5	Tedbîr-i Millî Bankası'na		106321	75		
9	Mobilyeye		11850	—		
8	Eshâm-ı tahvîlât		172540	—		
15	Tahsîl senedine		40500	—		
14	Fî'lik efendiye		10000	—		
16	Edhem Beğe		6650	—		
25	Sermed Efendi'ye		5000	—		
26	Emlâke		229655	—		
			3459061	50	3459061	50

Appendix 5

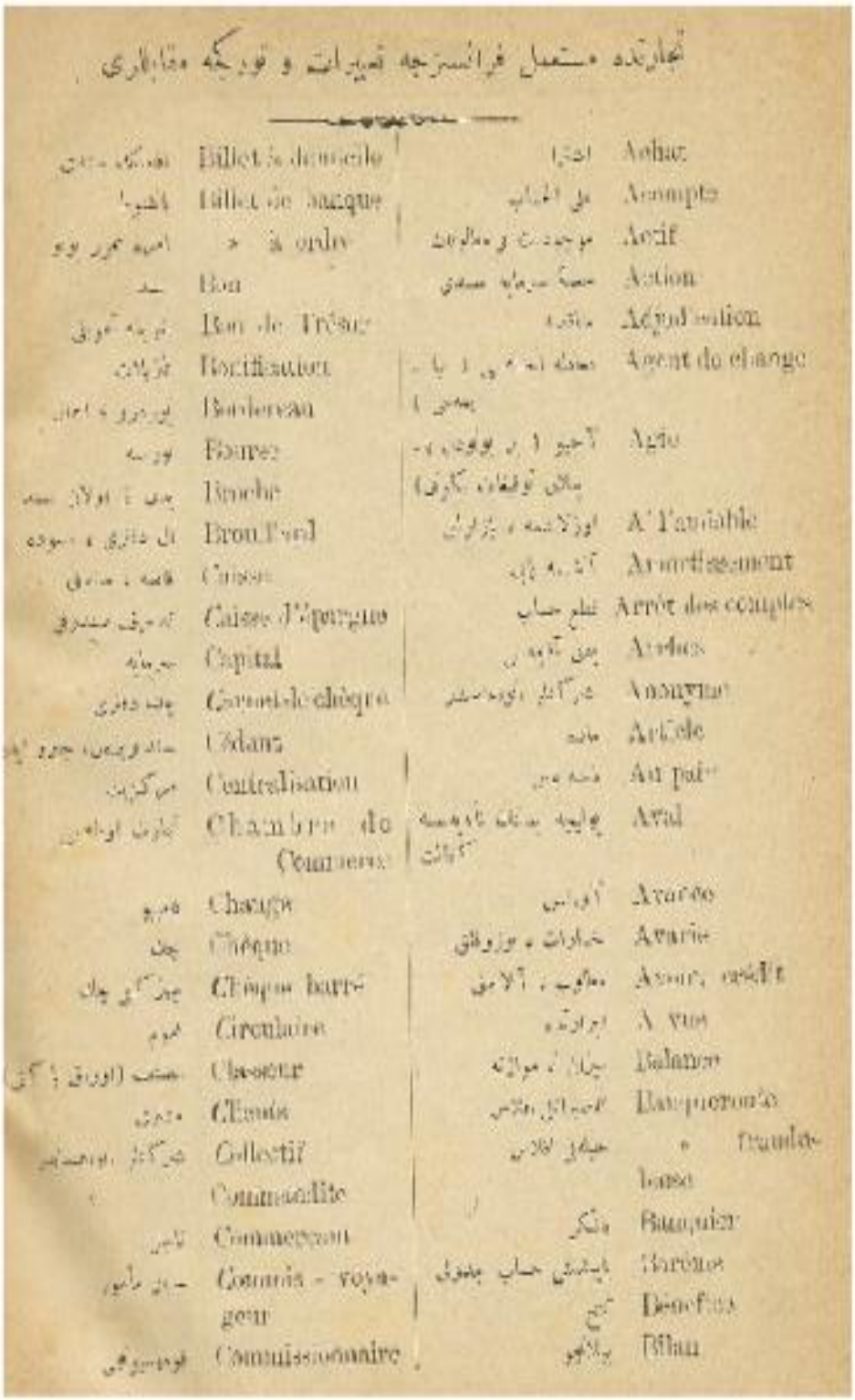
Cover of Accounting Textbook in the Armenian Language (1830)
(Source: Guvemli, 19. Yuzyilda Turkiye'de Muhasebe Ogretim Kitaplari -1997)



Appendix 6

First Page of the Glossary – Krikor Kömürcüyan (1926)

Merchant's Accounting Book



تجارتده مستعمل فرانسزجه تعبيرات و توركجه مقابلاري

Billet à domicile	Aehat
Billet de banque	Acompte
» à ordre	Actif
Bon	Action
Bon de Trésor	Adjudication
Bonification	Agent de change
Bordereau	
Bourse	Agio
Broche	
Bron Fond	A l'aviable
Caisse	Amortissement
Caisse d'Epargne	Arrêt des comptes
Capital	Arrhes
Carnet de chèque	Annonce
Cédant	Article
Centralisation	Au pair
Chambre de Commerce	Avil
Charge	Avarce
Chèque	Avarie
Chèque barré	Avoir, crédit
Circulaire	A vue
Classeur	Balance
Clients	Banqueroute
Collectif	» frauduleuse
Commandite	
Commerçant	Banquier
Commis - voyageur	Bareme
Commissionnaire	Benefice
	Bilan

Appendix 7: List of the Accounting Related Textbooks Compiled by Oktay Güvemli (1997) by 1928
(They are currently existing at the different libraries of Turkey)

(**Bold: Auditing;** *Italic: Analyses*)

In the Ottoman Empire		
Author	**Year published**	**Title**
Ahmed Tevhid	1855 (Hicri 1270)	Nuhbet-ül Hesab
Ahmed Tevhid	1861 (Hicri 1276)	Muhasebe Efendiler İle Mal Katiplerinin Vezaifini Şamil Talimatı
Ahmed Tevhid	Not known	Muhasebe-i Hususilerde Tutulacak Kayıt ve Defter-i Hesabiye Hakkındaki Talimatnamenin Vilayet B,tçelerinin Sureti Tertib ve Tanzimine Dair Olan Mevaddını Muaddil Talimatname
Mehmet Tevfik	1862 (Hicri 1287)	Teshil-i Hesap
Mehmed Münir	1862 (Hicri 1287)	Fenn-i Defteri
Mehmed Münir	1863 (Hicri 1288)	Varidat-ı Oşiyesinin Müzayede ve İhale ve İdaresi Hakkında Talimat-ı Cedide
Yusuf Paşa	1873 (Hicri 1298)	Kavaid-i İlmi Hesap
Mehmed Nuri Şemsettin	1867 (Hicri 1282)	**Murakabe Risalesi**
Ziya Bey	1883 (Hicri 1298)	Fenn-i Usulü Defteri
Asaf Süleyman	1885 (Hicri 1300)	Muhtasar Usulü Defteri
Ahmed Ragıp	1880 (Hicri 1295)	Hesab-ı Muamelat
Hasan Tahsin	1884 (Hicri 1299)	Yeni Usul Defteri
Muhiddin	1885 (Hicri 1300)	Usul-ü Defter-i Cedid
Süleyman Asaf	1885 (Hicri 1300)	Mülkiye ve Askeriye İçin Nazari ve Ameli Muhasebe
Süleyman Asaf	1886 (Hicri 1301)	Muhasebe-i Bahriye Dairesinin Heyet-i Vezaifi Dahiliyesi
Şeyh M Nuri B Seyyid Hüseyin	1886 (Hicri 1301)	**Risale-i Murakabe**
Aram Margosyan	1888 (Hicri 1303)	Muhasebe-i Ziraiye
Aram Margosyan	1889 (Hicri 1304)	*Hesab-ı Tahlili*
F.I.C. (Translated by Ali Nihad Abdülaziz)	1895 (Hicri 1310)	Hazine-i Tedrisat Talimi Hesap-Kısım: 2
F.I.C. (Translated by M.Ali Nazıma)	1892 (Hicri 1317)	Ameli ve Nazari Yeni Usul

Appendix 7 (Cont.): List of the Accounting Related Textbooks Compiled by Oktay Güvemli (1997) by 1928 (They are currently existing at the different libraries of Turkey)

(**Bold: Auditing;** *Italic: Analyses*)

In the Ottoman Empire		
Author	**Year published**	**Title**
Ziyaeddin	1899 (Hicri 1314)	Fenni Defteri
Mehmet Celal	1899 (Hicri 1314)	Hesab-ı Ameli
Ali Nazıma	1900 (Hicri 1315)	Muhtasar Yeni Hesap
William James (Translated by İsmail Faik)	1905 (Hicri 1320)	Mükemmel Hesap
Bahri Arif	1907 (Hicri 1322)	Ameli ve Nazari Fenni Defter-i Kaydı Muzaaf
Mehmet Halit	1907 (Hicri 1322)	Nizam-ı Tatbik, Cilt: 1-2
Nail Beyefendi	1909 (Hicri 1324)	Fenn-i Usul-ü Maliye
Hüseyin Hıfzı	1912 (Hicri 1327)	Muhtasar Hesap, 2.baskı
Hüseyin Hıfzı	1912 (Hicri 1327)	Devr-i Muhasebe Defterinin Suret-i Tanzimine Dair Talimat
Hamid	1912 (Hicri 1327)	Muhtasar-ı Usul-ü Muhasebe, Kısm-ı Sani, Fenni Defteri
Faruk	1913 (Hicri 1328)	Hesab-ı Ameli-i Mali
Berberyan, Surpe, Nevrez	1913 (Hicri 1328)	Usul,i Muhasebe,i Umumiye,
Berberyan, Surpe, Nevrez	1913 (Hicri 1328)	Usul,i Muhasebe,i Umumiye-Kanun-i Şerhi
Hüseyin Hıfzı	1913 (Hicri 1328)	Sualli Cevaplı Defter Tutmak Usulü
Ali Hafız	1909 (Hicri 1324)	Yeni Usulü teshili Hesap
İstepan Arapyan	1915 (Hicri 1330)	Usul-ü Defter-i Mali, 2. Cilt
İstepan Arapyan	1916 (Hicri 1331)	Usul Defteri
Ahmed Nazif	1916 (Hicri 1331)	Usul-i Aşair
Serkiz Nihad	1917 (Hicri 1332)	Defter-i Kebirli Yevmiye
Serkiz Nihad	1917 (Hicri 1332)	Muhtasar Mhaberat-ı Tüccariye
Serkiz Nihad	1919 (Hicri 1334)	Usul-ü Muhasebe-i Maliye, yahut Banka, Borsa, Sigorta Muhasebeleri
Kirkor Kömürcüyan	1920	Hesab-ı Ticari ve Mali
İstepan Arapyan	1922	Yeni Usul-i Defteri

Appendix 7 (Cont.): List of the Accounting Related Textbooks Compiled by Oktay Güvemli (1997) by 1928 (They are currently existing at the different libraries of Turkey)

(**Bold: Auditing;** *Italic: Analyses*)

In the Republic of Turkey		
Serkiz Nihad	1924	Muhasebe-i Tüccariye Mehadisi
İbrahim Fazıl	1924	Amili Usul-ü Muhasebe-i Ziraiye ve Çiftlik Muhasebesi
Serkiz Nihad	1926	Sigorta Muhasebesi
Serkiz Nihad	1926	Nazari ve Ameli Muhasebe
Kirkor Kömürcüyan	1926	Esnafın Hazır Defteri
Kirkor Kömürcüyan	1927	Yeni Hesabı Ticari
Şakir Mehmet	1928	Türk Usulü Yeni Defter

How Do Investors Process Pay Equity Disclosures? Experimental Evidence with Pay Ratio Disclosures from Section 953(b) of the Dodd-Frank Act

Abstract: The pay ratio disclosure requirement established under Section 953(b) the Dodd-Frank Act requires most publicly traded companies to report the compensation of their Principal Executive Officer (PEO), the compensation of a median employee, and the ratio between the two. The primary contention of this requirement was if and how the information would be useful to investors. A behavioral experiment is conducted to examine how investors might use the pay ratio disclosure in making their evaluations of companies. This study uses a 3 × 3 design with the manipulation of PEO and median employee compensation with five possible resulting pay ratios. It is found that investors process the information from the pay ratio disclosure through a heuristic that jointly considers fairness and the failure of labor markets. These judgements are then used by investors to evaluate the attractiveness of their investment, which in turn influences voting behavior in say on pay votes.

Data Availability: Contact the author

Keywords: Compensation disclosures, Dodd-Frank Act, judgement and decision making, pay ratio disclosure

I. INTROCUCTION

The pay ratio disclosure mandated by Section 953(b) of the Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank Act) provides new information on rank-and-file employee compensation and how it compares to executive compensation. Investors should be able to use this new disclosure for its stated purpose of assessing executive compensation (Securities and Exchange Commission 2015), although they should also be concerned about the newly-available information on rank-and-file employee compensation and the indirect impact on the company based on the reactions of other stakeholders (Fitz-enz 1997; Mohan et al. 2015). It is not known how investors and other stakeholders will perceive and analyze the disclosed compensation of the Principal Executive Officer (PEO) and the median employee or how those perceptions will affect investor judgements regarding the value of firms and their investment decisions. Therefore, it is important to evaluate the context of the compensation-setting process to promulgate an understanding of the economic and social theories through which investors and other stakeholders will contextualize and evaluate the information that can be derived from the pay ratio disclosure.

Investors have a direct interest in the compensation of both executives and rank-and-file employees as those compensation policies relate to and reflect the performance of the company. An investment may also be impacted by the reaction of employees (Akerlof and Yellen 1990; Ramlall 2004) and consumers to these compensation policies (Castaldo et al. 2009; Mohan et al. 2015). Whereas the assessment of executive compensation by investors based on prior disclosures has been relatively well examined (Agarwal 1981; Devers et al. 2007; Garen 1994; O'Reilly et al. 1988; Werner and Ward 2004), the assessment of rank-and-file employee compensation has been limited due to a lack of disclosed information (Lajili and Zéghal 2005).

The Dodd-Frank pay ratio disclosure further contextualizes both executive compensation and median compensation in relation to one another, and this may further affect how investors process both elements of information.

Investors should be able to use the new Dodd-Frank pay ratio disclosure to assess the adequacy or excess of executive compensation, and they should use those assessments in making investment decisions and in say-on-pay proxy votes regarding executive compensation. The contracting of executive compensation should resolve the agency problem between shareholders and management (Gomez-Mejia and Wiseman 1997; Jensen and Meckling 1976). However, it is unclear whether the market for executive talent can achieve optimal contracting (Edmans and Gabaix 2009; O'Reilly et al. 1988) or if the existence of managerial power in the compensation setting process enables the extraction of rents (Bebchuk and Fried 2005; Bebchuk and Fried 2003; Gomez-Mejia et al. 1987; Redling 1981; Sridharan 1996). The extraction of rents by executives would reduce the value of the company for investors, and so they should be more likely to vote against a say on pay proxy vote and be less willing to invest in a company if they perceive that executives are excessively compensated (Kelly and Seow 2016; Kimbro and Xu 2016).

Investors should also be concerned with the rank-and-file level employee compensation data that has become available as a part of the Dodd-Frank pay ratio disclosure. The compensation of the median employee is information that employees can use as a point of comparison in assessing the equity of their own compensation (Adams 1965). If employees believe that they are not being equitably compensated, they may impose costs related to a lack of motivation or organizational commitment (Akerlof and Yellen 1990; Cowherd and Levine 1992; Ramlall 2004) or related to separation from the organization (Ballester et al. 2002; Ho et al.

2009; Wyatt and Frick 2010). The data on the compensation of the median employee may also be value-relevant to investors regarding a company's investment in human capital (Lajili and Zéghal 2006; Lajili and Zéghal 2005) or regarding risks associated with the labor markets in which a company operates (Frölich and Haile 2011).

This study uses an experimental methodology to assess the processes through which investors will develop perceptions of executive compensation and rank-and-file compensation from the Dodd-Frank pay ratio disclosure and how those perceptions will affect investor assessments of the firm and their propensity to vote against in say-on-pay votes for executive compensation. Investors may be able to directly interpret the economic value of compensation and how that signals the value of the employee to the firm, in which cash the information in the pay ratio disclosure should directly affect investor perceptions of the firm as an investment and how they vote in the say-on-pay. However, if investors lack this ability, it is posited that they will instead rely on a heuristic evaluation of the fairness of compensation and perceptions labor market failures to set reasonable compensation, which will then affect their perceptions of the firm and their votes in the say-on-pay.

This study provides evidence that the Dodd-Frank pay ratio disclosure is useful to investors in making assessments about companies and for deciding how to vote on executive compensation. These results suggest that investors use a heuristic process involving fairness and perceptions of labor market failures to process the pay ratio disclosure information into useful, company-level assessments. These assessments influence perceptions of the attractiveness of the company as an investment, and the perceived attractiveness drives investor decisions in say on pay votes. Collectively, the results suggest that investors are attempting to make rational economic choices with the limited information available.

II. THEORETICAL BACKGROUND AND LITERATURE REVIEW

Efficient Labor Markets and Compensation Disclosures as a Signal of Value

In an efficient and perfectly competitive labor markets, compensation would be determined by the marginal product of labor. Although discrepancies between supply and demand for labor would create temporary deviations from this efficient wage, it would stabilize over time as employees would change their willingness or ability to supply the requisite labor and companies would enter the market for that labor or find substitutes (Hyclak et al. 2004). Under these conditions, the compensation paid to an employee would signal the productivity of that employee and thus the value to the employee to the company. Although a perfectly efficient labor market would not allow economic rents by either employees or employers in the long run, deliberate strategies to employ relatively high- or low-cost labor or temporary deviations based on supply and demand could affect the future cash flows of firms and thus their valuation (Balkin and Gomez-Mejia 1990).

The ability of investors to price compensation and compensation strategies into securities prices also requires compensation information to be known. In an environment with perfect information, the compensation of all employees would be known, and thus the total product of labor and the average product of labor, or productivity, per employee could be derived. Even in an environment without perfect information, any compensation information that is revealed would signal the relative productivity of the employee in question if the labor market is still otherwise efficient. These signals would then be interpreted by investors and other stakeholders and imputed into the valuation of the company (Hyclak et al. 2004).

[Insert Figure 1 Here]

Executive Compensation, Optimal Contracting, and Executive Power

In an efficient market for executive talent, companies and potential executives would arrive at an arms-length, optimal contract that would incorporate all information and could not be unduly influenced by either party (Edmans and Gabaix 2009). However, the potential for management to affect the actions of the company towards their own ends and to affect the compensation-setting process by influencing boards of directors suggests the extraction of rents that reduce the value attributable to shareholders (Bebchuk and Fried 2003; Daily et al. 1998; Finkelstein and Hambrick 1989; Sridharan 1996). Prior literature suggests that these agency problems can be ameliorated with increased oversight on executive compensation through increased disclosures (Vafeas and Afxentiou 1998), better corporate governance structures (Core et al. 1999), institutional investors scrutiny (Hartzell and Starks 2003), and shareholder votes on executive compensation (Kimbro and Xu 2016).

Absent perfect information, compensation committees must rely on available information even if it does not perfectly reflect the value generated by the labor of the executive. One of the primary factors identified is firm size, which accounts for a large proportion of the cross-sectional variation in executive compensation, and it is argued that larger companies require more effort to manage and so necessitate greater compensation (Kostiuk 1990; Tosi et al. 2000). Surprisingly, firm performance explains relatively little of the variation in the compensation of executives (Agarwal 1981; Cordeiro and Veliyath 2003; O'Reilly et al. 1988), which suggests an inability to resolve the agency problem through compensation contracting (Gomez-Mejia et al. 1987; Tosi et al. 2000). However, other findings suggest that managerial attributes (Graham et al. 2012) and abilities (Custódio et al. 2013; Demerjian et al. 2012) are significant factors in the

determination of compensation, which suggests that compensation committees are considering inputs as well as outputs in the design of compensation contracts (Bushman et al. 1996; Finkelstein and Boyd 1998; Indjejikian 1999).

Compensation committees may also substitute market-based and other comparison information to benchmark executive compensation. Although peer comparisons can create normative expectations of executive compensation, the selection of a peer group is subjective and thus can be manipulated (Byrd et al. 1998; Gong et al. 2011). The hierarchies of a company's bureaucracy may also influence executive compensation as those at the higher levels will be expected to earn more than those at the lower levels (Hall and Liebman 1998).

Executives are expected to act as agents for shareholders of the company that they manage (Jensen and Meckling 1976) , but the power that executives hold provides the opportunity to affect their own compensation (Garen 1994; Grabke-Rundell and Gomez-Mejia 2002). If this occurs, executive compensation will deviate from optimal contracting, economic rents will accrue to the executives, and the value of the company remaining for shareholders will be lower (Bebchuk et al. 2002). These economic rents extracted by executives can take various forms depending on the circumstances of the business and the monitoring of boards and investors including higher total compensation, less compensation at risk (Bebchuk and Fried 2005), and alternative form of compensation such as perquisites (Yermack 2006).

The ability of executives to exert power to increase their own compensation requires that they are able to acquire such power and influence within an organization and that they are not prevented from doing so by monitoring functions (Ben-Amar and Zeghal 2011; Daily et al. 1998; Mallette et al. 1995). Evidence from Hill and Phan (1991) suggests that executives are able to achieve a compensation package that is less aligned with stock returns, and thus their own

performance, as their tenure increases. Executives that serve as chair of the board have a greater scope of power and have greater ability to influence the compensation setting process (Main et al. 1995), whereas boards with greater independence, particularly with the independence of the compensation committee, are better able to achieve efficient compensation contracts (Vafeas 2003; Sridharan 1996). Large institutional shareholders, who generally have a direct ability to communicate and influence the board of directors, are able to limit excessive executive compensation (Hartzell & Starks, 2003). Individual investors may now also participate in this monitoring function through say-on-pay proxy votes, and evidence suggests that they are effective in reducing the growth of executive compensation in future periods (Ertimur et al. 2013).

Rank-and-File Employee Compensation, Human Capital, and Labor Market Inefficiencies

The compensation of rank-and-file employees is less economically complex due to the significantly lower level of agency held by rank-and-file employee, although the social complexities of rank-and-file employees may be greater than for executives because rank-and-file employees are more heterogenous in their preferences and subject to greater bureaucratic influence. Human capital is similarly salient with regard to the compensation of rank-and file employees given that it is an observable way to assess the potential productivity of employees (Ballester et al. 2002). Assessing performance beyond verifiable measures of human capital generally requires the bureaucratic process of performance evaluation, which has the ability to align compensation outcomes with performance for rank-and-file employees (Widener 2006) but is also susceptible to bias and other inefficiencies (Fisher et al. 2005). Employees may also seek nonfinancial characteristics of employment such as convenience, status, and social

considerations and be willing to accept lower wages in exchange for that utility (Hodson 1989; Phelan 1994).

Rank-and-file employees generally lack the power to introduce inefficiencies into labor markets. However, employers may able to do so if they are part of an oligopsony within the labor market and can thereby affect the demand for labor. Under these circumstances, companies can differentiate themselves within the labor market by changing their compensation practices, and if the oligopsony power is significant, companies in a labor market could cause rank-and-file employees to accept lower wages if the alternative is unemployment (Bhaskar et al. 2002). The exercise of oligopsony power will be more effective in labor markets that require less human capital because they will have greater competition among employees due to a larger, less-differentiated supply of labor (Bhaskar and To 2003).

Rank-and-file employees may also impose direct and indirect costs on employers by withholding effort and productivity, imposing frictional costs due to transitions, and engaging in direct actions against the employer. Equity theory (Adams 1965) and by extension the fair-wage effort hypothesis (Akerlof and Yellen 1990) posits that employees will withhold effort to achieve equity if they perceive their compensation to be inadequate. Withheld effort based on perceptions of equity can have real economic consequences for companies such as lower product quality (Cowherd and Levine 1992). Employees may also choose to leave a firm based on their perceptions of compensation (Ho et al. 2009), which imposes costs due to the loss of human capital and institutional knowledge and hiring costs to replace those employees (Bothma and Roodt 2013; Fitz-enz 1997). Additionally, Employees are more likely to engage in direct action against their employer if they perceive their compensation to be unfair (Chen and Sandino 2012).

Market Inefficiencies and the Information Content of Compensation Disclosures

The existence of inefficiencies in the labor market either through contracting for compensation or distortions of supply and demand would affect compensation such that it no longer correctly signals the value of the employee to the company (Boeri and Van Ours 2013). In executive labor markets, these deviations of compensation from productive value would also be relevant information to shareholders as any compensation paid to executives in excess of the value they provide to the company would be an economic rent accruing to executives that would reduce the value of the company compared to the economically efficient outcome (Bebchuk and Fried 2003). In labor markets for rank-and-file employees, deviations in compensation could represent additional costs due to a low supply of labor, which would reduce the value of the firm for investors (Wadongo and Abdel-Kader 2014), or represent structural choices in low-wage employment, which would increase the value of the firm and introduce additional risk (Frölich and Haile 2011).

Neither the value signal nor the noise can be precisely known, and so investors or other stakeholders must estimate them if the productive value of the employee is to be used in making decisions about the company. It should be expected that users of the compensation disclosures will use other available information that could differentiate the value signal from the noise (Connelly et al. 2011). The most available and relevant information would be comparative compensation information from other executives or employees within the same or similar labor markets. This market-based comparative information would be subject to similar market factors and thus contain much of the same noise relative to the signal. However, a breadth of such information would represent the market norms for compensation against which the compensation of a focal employee can be evaluated (Martin 2015). Considered in this context, users of a

compensation disclosure must estimate and assign a value for the productive value of the employees in question and any noise, and so this estimation will always be made with some degree of error.

This also raises the question regarding how investors and other stakeholders attempting to use compensation disclosures will process that information. Investors are unlikely to have the specialized knowledge of labor markets necessary to evaluate the effects of labor market inefficiencies and market norms effectively (Lajili and Zéghal 2005), and so it is expected that they will use heuristics to incorporate information from compensation disclosures into their decision making processes (Baker and Ricciardi 2014). The Dodd-Frank pay ratio disclosure provides the framing for this heuristic in the form of comparative analysis of compensation and the resulting perceptions of the functioning of labor markets and fairness of compensation.

[Insert Figure 2 Here]

Labor Markets and Fairness

Assessments of fairness resulting from labor markets should be consistent with the established dimensions of equity used in the assessment of compensation, external equity, internal equity, individual equity, and procedural equity. External equity is concerned with the proper outcome for an individual as compared to others in similar circumstances in other organizations. Internal equity is concerned with the relative outcomes of individuals in the same organization considering their positions and contributions. Individual equity is concerned with the relative inputs of individuals as compared to their relative outputs, particularly at the same

level or function within an organization. Procedural equity is concerned with the equitable treatment of individuals during the compensation setting process (Terpstra & Honoree, 2003).

Considering procedural fairness first, an understanding of the proper functioning of labor markets to achieve equitable outcomes requires the consideration of referent groups and the criteria to establish equitable outcomes or processes (Greenberg and Colquitt 2005). This conceptualization of procedural fairness holds that outcomes will be equitable if the organization treats individuals consistently and arrives at outcomes consistent with generally established standards (Price 1997). In this context, that would indicate that the compensation observed for other similar labor-market participants is fair if arrived at through an optimal contracting process. Therefore, the observation of a labor market outcome considered to be unfair may also be perceived as a failure of the labor market.

The external and individual fairness of executive compensation can be understood by considering the compensation-setting process and the power of executives to influence their compensation. A dearth of compensation should be interpreted as a choice of the executive considering that executive power effectively eliminates labor market mechanisms that would provide incentives for executives to accept compensation below an equity level (Levine 1993). This suggests that below-market compensation for executives should not have an adverse impact on the company due to executive labor concerns and that it should not be considered unfair due to the agency of the executive in accepting that compensation. An excess of compensation would indicate a lack of optimal contracting in some capacity and could indicate a labor market failure and unfair outcome (Bebchuk et al. 2002; Bebchuk and Fried 2005; Faulkender et al. 2010).

For rank-and-file employees, the external and individual fairness of compensation can be understood by considering that rank-and-file employees are not expected to have influence over

the functioning of labor markets, an excess of compensation for the median employee should be interpreted as either a market mechanism to allocate a scarce labor supply (Hyclak et al., 2004) or a strategy to maintain an effective workforce with lower transition costs (Chen and Sandino 2012; Fitz-enz 1997). In these cases, it would still be economically inefficient for a company to hire or retain employees under these conditions, and so it should be expected that these conditions satisfy perceptions of fairness. A dearth of compensation could be interpreted as a failure to sufficiently invest in human capital (Ballester et al. 2002; Lajili and Zéghal 2005; Wyatt and Frick 2010), as a risk factor related to labor force instability and transition costs (Frölich and Haile 2011), or as a distortion of labor markets due to oligopsony power held by large companies (Bhaskar and To 2003). Although these cases would all result in lower rank-and-file employee compensation, they may be perceived as simple unfairness relative to the market but also in some cases as the failure of the market to establish appropriate compensation.

Considerations of the internal fairness of compensation of can be useful to investors in assessing the labor markets and resulting fairness of both executive and rank-and-file employee compensation. Some level of dispersion within an organization is considered desirable to motivate employees to achieve greater productivity for a greater compensation (O'Reilly et al. 1988). However, employees who feel their compensation is not equitable based on pay dispersion may withhold effort and productivity or cause additional transition cost through turnover (Fitz-enz 1997). Customers have also been observed to decrease their purchasing intentions and have worse affective perceptions for companies with greater pay ratios (Mohan et al. 2015). Therefore, we should expect investors to consider how the internal fairness of compensation relative to an appropriate level of pay dispersion on a theoretical level and as observed by other market participants.

III. HYPOTHESIS DEVELOPMENT

Investors, employees, and other stakeholders can use the Dodd-Frank pay ratio disclosure to assess executive and rank-and-file employee compensation. As put forth by the SEC in the rules for the pay ratio disclosure, investors are expected to use the Dodd-Frank pay ratio disclosure primarily to assess executive compensation. However, that assessment of executive compensation is now in a context where rank-and-file compensation and pay dispersion are salient. Therefore, it is important to determine if investors will be able to evaluate compensation directly as a signal for productive value or if they will need to apply a heuristic related to labor market functions and fairness in order to form judgements and make decisions regarding the company and executive compensation. Early evidence from Kelly and Seow (2016) suggests that investors do assess pay equity in response to the pay ratio disclosure, and that those assessments are negatively related to executive compensation and positively related to median employee compensation.

[Insert Figure 3 Here]

The Signal of the Value of Employee Labor and Compensation Disclosures

Investors, employees, and other stakeholders can make use of compensation as a signal of the productive value of employees to the company. Under efficient labor market conditions, the compensation setting process should arrive at compensation levels that reflect the value that employees generate for companies (Hyclak et al. 2004). The possible inefficiencies in labor

markets have the potential to affect compensation so that it no longer fully signals the productive value of employees (Bebchuk and Fried 2005). Regardless, investors, employees, and other stakeholders should be able to derive some information about the productive value of employees from the compensation disclosed.

In the case of executive compensation, efficient contracting would suggest that the highest paid executives produce the greatest value for their companies. However, the inefficiencies that can exist in the executive labor market due to managerial power allow for a disconnection between levels of compensation and the value of an executive's labor. Despite these potential inefficiencies, the level of executive compensation can still serve as a signal of the value of an executive's labor given that some part of the compensation should be tied to actual value of productivity even though there may be additional compensation acting as upward noise with regard to the signal of value (Wyatt and Frick 2010).

H1a: The compensation of the Principal Executive Officer (PEO) is positively associated with the attractiveness of the investment.

In the case of rank-and-file employee compensation, an efficient labor market would imply that employees are paid consistent with the marginal product of their labor. However, inconsistencies in the labor market such as a shortage of skilled labor or oligopsony power of employers can cause the prevailing wage to deviate from the marginal product of labor. In such cases, the market wage for labor will deviate from the productive value of that labor. Despite the possibilities of market disequilibrium or inefficiencies, the level of compensation of the median

employee can still serve as a signal of the actual value of productivity of rank-and-file employees despite the possible noise, both upward and downward.

H1b: The compensation of the Median Employee is positively associated with the attractiveness of the investment.

If investors interpret the compensation as a signal of the productivity of an executive, they should carry this interpretation over into the say on pay vote for executive compensation. This conceptualization holds for efficient labor markets in which executive compensation is entirely determined by executive talent and the supply and demand for executive labor. The eventual conclusion of this conceptualization of the executive labor market is that investors should support higher levels of executive compensation because they must be associated with higher levels of executive talent and productivity.

H1c: Support for the say on pay proxy vote is positively associated with the compensation of the Principal Executive Officer (PEO).

Heuristic Processing of Compensation Considering Labor Markets Failures and Fairness

The effects of the Dodd-Frank pay ratio disclosure on investor perceptions of and decisions about the company could be made by evaluating perceptions of labor markets and fairness. Development of measures considering perceptions of external, individual, internal, and procedural fairness revealed that investors process external, individual, and internal fairness as a single construct of overall fairness of the compensation of a focal company and that they process

procedural fairness and internal fairness with a separate factor relative to the failure of labor markets to establish or maintain appropriate compensation (Craven 2020). These perceptions comprise a heuristic that allows investors to more easily process the complex interactions and context of compensation for use in decision making (Baker and Ricciardi 2014).

Labor Market Failures

A perfectly efficient functioning of labor markets should result in optimal contracting for compensation. However, limited information and the ability of certain parties to affect the market can move compensation away from optimal levels. It is possible to investors and other stakeholders to make judgements about labor markets from the resulting compensation based on the expected mechanisms of labor markets to derive that compensation. Investors may then form other assessments of that compensation, such as its overall fairness, based on their perceptions of the process to set that compensation.

The association between executive compensation and the failure of labor markets can be understood in terms of perceptions of executive power and social comparisons of executive compensation. Primarily, this relates to the ability of executives to exert power on their companies to achieve higher compensation that is associated with the violation of labor market mechanisms and norms for executive compensation. However, below-market compensation for executives should not necessarily be interpreted as a market failure because the existence of executive power gives executives the choice to accept lower compensation. Social comparisons and improper benchmarking of executive compensation is another form of market failure that will raise executive compensation absent productive value of the executive. Therefore, it should

be expected that PEO compensation should have a linear and positive relationship with perceptions of labor market failures.

H2a: The compensation level of the PEO is positively associated with the perceptions of labor market failure.

The association between the compensation of the median employee and labor market failures will instead rely primarily on social comparisons that enable evaluations of individual equity given that an individual rank-and-file employees lack the ability to affect labor markets. Given that investors should have some conception of rank-and-file compensation from their own compensation and economic indicators of compensation, they should use these benchmarks to make comparisons to the reported compensation of the median employee, and they will develop perceptions of the adequacy of that compensation in terms of economic value. This means that assessments of median compensation in this context should focus on low levels of median compensation. Although higher levels of compensation for median employees should be attributable to market forces, there may be spillover effects of high median compensation in comparison to executive compensation to suggest higher labor market failures for the organization. Therefore, it should be expected that the compensation of the median employee should have a linear and negative relationship with perceived labor market failures.

H2b: The compensation level of the Median Employee is positively associated with the perceptions of labor market failure.

Finally, perceptions of optimal pay dispersion suggest that individuals hold a much lower optimal pay dispersion than what is observed among publicly-traded companies (Kiatpongsan and Norton 2014). This disconnect suggests that investors are likely to hold perceptions of a labor market failure based on this ideal pay ratio compared to the observed ratio while holding perceptions of labor market failures based on the market-standard ratio compared to the observed ratio. Considering that expectations of pay ratios tend toward the lowest levels of the observed pay ratios, it is expected that the pay ratio will be positively and linearly associated with labor market failures.

H2b: The pay ratio is positively associated with the perceptions of labor market failure.

Fairness

Perceptions of fairness incorporates the dimensions of the external, individual, and internal equity of compensation relative to comparative information from within the company and to labor market standards. In making these comparisons, investors are developing a holistic view of the compensation practices of a company from the individual components of the Dodd-Frank pay ratio disclosure. Most prior conceptualizations of organizational justice measure it based on the perceived justice to an individual (Greenberg and Colquitt 2005; Price 1997), and so it is important to note that investors are evaluating fairness at the organizational level.

Considering fairness and executive compensation requires the reconciliation of the market-standard compensation and executive power. Compensation that is set at the market-standard rate should be associated with a middling to higher perceptions of fairness as the compensation is consistent with the market norms. Compensation that is higher than market-

standard may be an outlier of productivity, but it is more likely to be the result of executive power creating inefficient compensation, which should be associated with lower levels of perceived fairness. Finally, lower levels of executive compensation are not expected to be associated with lower levels of fairness given that this deviation should be a choice by the executive. Therefore, executive compensation is expected to have a linear and negative association with fairness.

H3a: The compensation level of the PEO is negatively associated with the perceptions of fairness.

The consideration of the effect of median compensation on fairness is less complex because rank-and-file employees should not be expected to have the power to affect labor markets. Further, companies that may affect labor markets through oligopsony power are affecting the entire market and not affecting particular employees. Therefore, the effect of median compensation on fairness should be based on the deviation from the market-standard compensation with any compensation at or above the market standard being considered more fair. The compensation of the median employee is expected to have a linear and positive association with fairness.

H3b: The compensation level of the median employee is negatively associated with the perceptions of fairness.

The pay ratio provides another basis of comparison for the dispersion of compensation within an organization, and it has the potential to reveal a compensation strategy for the firm as a whole. A relatively low ratio in combination with high executive and median compensation may suggest that a firm pays above-market compensation to all of its employees to enjoy the lower turnover and higher productivity (Chen and Sandino 2012; Fitz-enz 1997). Conversely, a comparatively high ratio can only be achieved with some combination of lower median compensation and higher executive compensation, and this may exacerbate perceptions of inequity. Therefore, it is expected that the pay ratio will be negatively associated with perceptions of fairness.

H3c: The pay ratio is negatively associated with the perceptions of fairness.

Perceptions of fairness capturing the external, individual, and internal fairness of compensation within a company must also consider the procedural fairness of compensation captured through the perceptions of labor market failures. If a labor market failure occurs, the resulting compensation will deviate from the established norm, and will therefore guarantee a deviation from external fairness. Although it is possible to meet individual and internal fairness under these conditions, it is not likely given that the possibilities for market failures will tend to result in excess executive compensation and lower median compensation. Therefore, the perception of labor market failures is negatively associated with perceptions of fairness.

H3d: Perceptions of labor market failures is negatively associated with the perceptions of fairness.

Attractiveness of the Investment

The attractiveness of the investment, as used in the context of this study, captures both the economic and affective value of the investment to the investor. If investors are using perceptions of labor market failures and fairness in compensation as an information processing heuristic, those perceptions should affect their perceptions of the attractiveness of their investment. These perceptions of the value of and expectation for the firm should then influence the judgements and investment decisions of the rational investor.

Considering executives, compensation levels that exceed the market standards are often the result of the extraction of economic rents by the executive through the exercise of executive power (Bebchuk et al. 2002). These economic rents are a deadweight loss to the company and will therefore lower its economic value. Considering rank-and-file employees, low compensation relative to the market should signal the employment of lower-productivity labor (Akerlof and Yellen 1990) or operations in higher-risk jurisdictions with lower labor costs (Frölich and Haile 2011), both of which connote lower economic value for the company. These explanations suggest that lower perceptions of fairness will be associated with lower perceptions of the attractiveness of the investment.

H4a: Perceptions of the attractiveness of the investment are positively associated with perceptions fairness.

The perception of market failures is expected to affect the attractiveness of the investment related to the expected economic effects of those market failures. With regard to

executive compensation, it is expected that high executive compensation will be perceived as a market failure when the exercise of power by executives allows them to defy market standards in order to extract economic rents (Bebchuk et al. 2002). For median employee compensation, labor market failures suggests lower productivity and higher employee turnover, and this should be expected to decrease value from productivity and increase risk. Finally, individuals tend to hold a much lower perception of the optimal pay dispersion than the market allows, and so higher pay ratios may also help to signal a market failure of executive and/or median compensation. All of these factors are complementary such that higher levels of perceived market failures should be associated with a lower perceived attractiveness of the investment.

H4b: Perceptions of the attractiveness of the investment are negatively associated with perceptions of market failures.

Say on Pay Votes and Perceptions of the Company

The Dodd-Frank Act provides for say on pay votes whereby investors vote for executive compensation on an advisory basis. The result of these say on pay votes reveal investor perceptions of executive compensation, both in terms of the value of the labor of the PEO and the equity of that compensation. Evidence suggests that companies respond to significant negative votes by investors (Ertimur et al. 2013; Kimbro and Xu 2016).

Affective effects on Opposition to the Say on Pay Vote

Although these say on pay votes are primarily concerned with executive compensation, the votes of investors may also take into account the compensation practices of the organization

as a whole. The perceived market failures and perceived fairness of compensation as a heuristic include an affective view of the company and its compensation practices, and this affective evaluation may be directly applied to the say on pay vote. Investors may wish to punish executives for excessive compensation notwithstanding any opposition related to lower perceived attractiveness of the investment.

H5a: Opposition to the say on pay proxy vote is negatively associated with perceptions of fairness.

H5b: Opposition to the say on pay proxy vote is positively associated with perceptions of labor market failures.

Attractiveness of the Investment and Say on Pay Votes

The say on pay proxy vote is the opportunity for shareholders to support or oppose the compensation of executives. In addition to the comparisons of compensation within the company and to industry averages to evaluate the equity of executive compensation, investors have the opportunity to evaluate the internal equity of executive compensation relative their return on investment or their valuation of the investment, as represented by the attractiveness of the investment. This is consistent with the use of pay-for-performance compensation to resolve the agency problem for executive compensation (Cordeiro and Veliyath 2003; Core et al. 1999; Fatemi et al. 2003). Therefore, investors should support higher levels of executive compensation when their valuation of the company is higher, and so it is expected that the attractiveness of the investment will be negatively associated with opposition to the say on pay proxy vote.

H5c: Opposition to the say on pay proxy vote is negatively associated with the attractiveness of the investment.

IV. EXPERIMENTAL DESIGN AND VARIABLES

Experimental Design and Manipulations

The interpretations and effect of the pay ratio disclosure are tested with an experiment that different levels of PEO compensation, median employee compensation, and the pay ratio. The PEO compensation and the compensation for the median employee are each manipulated with three levels: above average, average, and below average relative to the industry. This results in a fully crossed 3 × 3 between subjects experimental design. Although the pay ratio is also included as part of the disclosure, it is derived directly from the PEO compensation and the compensation of the median employee, and so it cannot be manipulated independently. Therefore, the levels of PEO compensation and median employee compensation have been chosen to create only five different pay ratios from the nine possible combinations of PEO and median employee compensation.

[Insert Figure 4 Here]

Compensation levels for the PEO compensation are chosen to approximate the 25th, 50th, and 75th percentiles of the observations for the business services industry[1] to provide significant variation without the possibility of choosing an outlier. These were also scaled to ensure that

[1] The business service industry was chosen because (1) there was a suitable range of PEO and median employee compensation in the observed data create reasonable and effective manipulations, (2) individuals without specialized knowledge should understand what the company does, and (3) the industry lacks the specificity for individuals to have preconceived ideas about the employees or their compensation.

there would be the same factor between the below-average condition and the average condition as between the average condition and the above-average condition. To ensure the consistency of the pay ratio among conditions, the manipulations of the compensation for the median employee were also required to maintain the same factor between conditions. The chosen compensation levels for the median employee are consistent with the 5th, 50th, and 95th percentiles of the actual observations for median compensation in the business services industry. However, this discrepancy is likely due to the differences in the distributions for PEO compensation and median employee compensation as the five possible pay ratios derived from the manipulations are consistent with the 5th, 25th, 50th, 75th, and 95th percentiles for pay ratios in the industry.

Experimental Procedures and Materials

Participants responded to a survey delivered electronically via Qualtrics in order to assess reactions to the pay ratio disclosure and to items designed to measure the mediator variables. Participants were randomly assigned to one of the nine possible conditions. All outcomes and scales are measured for all participants. The complete survey instrument and procedures can be found in Appendix A.

Participants are instructed that they have invested $5,000 in the putative advertising company in the business services industry and that the company has average but consistent financial performance for its industry. The exposition on financial performance is meant to be innocuous and to discourage participants from considering favorable or adverse financial results. Participants are told that they have recently received a proxy statement and reminded of their ability to vote on shareholder proposals, including the say on pay vote for executive compensation.

Regardless of which manipulations they receive, participants are provided with the same text of the pay ratio disclosure with the exception that the values for PEO compensation, median employee compensation, and the pay ratio are replaced with the appropriate manipulations. The content of the pay ratio disclosure is based on actual disclosures from the business services industry. A simple and relatively short pay ratio disclosure was chosen for this instrument to avoid complications that would arise from justifications or explanations of either of the compensation elements or the pay ratio. Immediately after but simultaneously with the pay ratio disclosure, participants are told that with some research they are able to find the averages for each manipulation and the pay ratio for the industry. The average values provided are based on the average conditions for the manipulation rounded down to a relevant multiple resulting in averages of $7,250,000 for PEO compensation, $47,500 for median employee compensation, and 150 for the pay ratio.

Participants are then be presented with and respond to the measures of the primary variables of interest and then the mediator variables. The first variable of interest presented is the say on pay proxy vote, which requires participants to vote “for”, “against”, or “abstain” on the advisory shareholder vote regarding executive compensation. An alternate assessment of this variable is also included in which the participants are asked to rate their approval of the say on pay proxy vote on a scale of “strongly oppose” to “strongly support”. The text for this section was compiled from actual proxy statements but simplified for the purposes of the instrument. Following the say on pay proxy vote, participants respond to the scale items for the attractiveness of the investment scale that is comingled and ordered randomly with the fairness and labor market failures scales.

All participants end the survey by responding to manipulation checks and demographic questions. The manipulation check questions ask participants to evaluate the relative difference between the provided information and the industry averages, which were held constant across all conditions. Basic demographic information is collected along with information on participant's education and experience regarding investing.

Participants

A pilot test of the survey instrument was conducted with 50 participants through Amazon's Mechanical Turk (MTurk) system. These results are used to test the measurement scales and the manipulation of the variables of interest. After validating and finalizing the scales, the final survey instrument was distributed to another 200 participants through MTurk. The use of participants drawn from MTurk to proxy for unsophisticated investors is reasonable considering the broad nature of the disclosure and consistent with prior research (Brandon et al., 2014; Buchheit et al., 2018; Mason & Suri, 2012).

Manipulation of the Independent Variables

PEO Compensation

The three levels of PEO compensation used are $1,510,215, $7,381,932, and $36,082,882, which represent the below-average, average, and above-average conditions respectively. These values correspond to the 25^{th}, 50^{th}, and 75^{th} percentiles of the observed data for the business services industry and were chosen to maximize the external validity of the study. Levels of executive compensation for the business services industry tend to be clustered around the median but significant outliers are also present in the distribution. Thus, the interquartile

range as used in this study is broadly representative of a preponderance of the companies in the business services industry.

Median Employee Compensation

The three levels of median employee compensation used are $9,856, $48,176, and $235,485, which represent the below-average, average, and above-average conditions respectively. These values correspond to the 5^{th}, 50^{th}, and 95^{th} percentiles of the observed data for the business services industry. Considerations in selecting these values were balanced between maintaining the consistency necessary to limit the number of possible pay ratios while still ensuring the external validity of the study. Levels of compensation for median employees within the business services industry are more evenly distributed across the spectrum of values, and so this range is still broadly representative of a preponderance of the companies in the business services industry.

Pay Ratio

The pay ratio disclosure of the Dodd-Frank pay ratio disclosure is a linear combination of the PEO compensation and the compensation of the median employee. Therefore, a manipulation in one or both of those factors results in a corresponding change in the pay ratio. Although this would suggest that the effect of the pay ratio disclosure is the same as the interaction of PEO compensation and the compensation of the median employee, the presentation of the pay ratio as a part of the pay ratio disclosure allows for the magnitude and other elements of that presentation to affect the reactions of investors and employees independently of the PEO and median employee compensation (Nelson & Rupar, 2014).

The manipulation of PEO compensation and median compensation by a consistent factor creates values for the pay ratio that overlap across multiple conditions such that among the nine conditions from the manipulations there are only five possible pay ratios. Each level of each of the manipulated factors is associated with three different ratios, which should ensure a complete evaluation of the pay ratio separable from the PEO and median compensation elements. These five possible pay ratios derived from the manipulations are 6, 31, 153, 749, and 3,661, and they are consistent with the 5th, 25th, 50th, 75th, and 95th percentiles for pay ratios in the industry. The highest and lowest ratios are each associated with one combination of manipulations, the high and low ratios are each associated with two combinations of manipulations, and the average ratio is associated with three combinations of manipulations, as depicted in the pay ratio matrix in Figure 4.

Measurement of the Dependent Variables

Investor Reactions through Say on Pay Voting

Investor reactions regarding a say on pay voting decision is measured directly with a form designed to approximate actual proxy voting material for a shareholder vote in which investors must vote "for", "against", or "abstain". This presentation is consistent with the analysis, and this form of presentation maintains the external validity of the study. Although this design choice limits the responses to a categorical outcome (i.e. "for", "against", or "abstain"), the outcome of interest is votes "against" and the propensity for investors to vote against, which can be evaluated with this categorical measurement.

The categorical say on pay vote provides only limited information, and so an additional, more continuous measure of support or opposition to the measure is taken. This measure is taken

using a seven-point Likert scale ranging from “strongly oppose” to “strongly support”. The relatively continuous measure of support will allow for more nuanced analysis of investor sentiment regarding the say on pay vote and its antecedents. Additionally, the duality of the measurement may allow for the consideration of the level opposition required for investors to vote against the proposal per the categorical voting outcomes.

Investor Reactions Regarding the Attractiveness of Investment

The attractiveness of an investment represents broad sentiments of investors towards a company both financially and affectively. The development of this construct is the confluence of items designed to represent the perceived value of the investment and an investor’s willingness to invest in a company This multiple-item scale is more reliable than measures dependent on a single item and does not specify a specific time horizon unlike other scales of investment attractiveness (Kelly et al. 2012). It is capturing not only investing intentions but also affective commitment regarding recommendations and the alignment of values. This suggests that the attractiveness of the investment represents investors’ assessments of both financial and nonfinancial value of the a company, and therefore how attractive investment in that company is to the investor.

Measurement of Mediator Variables

Labor market failures and fairness are the mediating variables to evaluate how investors process the information from the pay ratio disclosure. However, these scales are derived from scale items designed to measure external equity of both the executive compensation and the compensation of the median employee and the internal equity of the comparison of the executive

and median compensation. The initial scales related to the perceived equity of compensation are developed using the perspectives on equity theory from Terpstra and Honoree (2003) regarding the external and internal dimensions of equity. Items for these scales are adapted from pay satisfaction questionnaire evaluated by Heneman and Schwab (1985), Scarpello and Huber (1988), and Summers and Hendrix (1991) and the organizational justice scales evaluated by Price (1997) and Greenberg and Colquitt (2005). All of these items were in the first person relative to the respondent in the original scales, and so they are adapted in these items to reflect a third-person perspective of the participant as investor and their evaluations of the company or the employee.

Fairness

Fairness reflects both the adherence of a company to labor market standards and the setting of compensation relative to productivity. These are evaluated for the PEO, the median employee, comparing the two, and at the company-wide level. The confluence of these items indicates that the aggregated responses are measuring fairness at the organizational level and that respondents as putative investors are developing perceptions of a characteristic of the putative company in which they have invested.

Labor Market Failures

Labor market failures reflects the feelings of investors regarding deviations from market standards and differences from individually-held conceptualizations of optimal or adequate compensation. These items are able to highlight perceptions of compensation that are separable from the normative compensation levels set by labor markets at all levels and reflect perceptions

of executive power in distorting the labor market This scale, therefore, succeeds in describing market failures differentiable from fairness because it captures sentiment where an individual believes that the a party has manipulated or circumvented labor markets to achieve non-market compensation.

V. RESULTS

Final Sample

The final sample includes the completed survey of 195 participants who were recruited on MTurk and completed the survey online through Qualtrics. The posting on MTurk generated 439 entries in the Qualtrics system that were fully resolved. Entries were screened for duplicate IP addresses to prevent ballot stuffing, and so 117 duplicate responses were removed. Of these 322 unique attempts, 95 individuals failed to answer the screening questions correctly, and 17 individuals failed the comprehension check. This left an initial sample of 200 complete survey responses. An analysis of the manipulation checks as detailed below revealed 5 participants as significant outliers in one or more of the manipulated conditions, and these participants are removed, which results in a final sample of 195 responses.

Demographics of the Participants in the Final Sample

The demographics of the final sample reveal a relatively broad sampling of individuals across gender, age, education, and income. Although a number of these deviate from the expected values of the general population, they are in line with the expected demographic makeup of MTurk workers (Mason and Suri 2012). This sample did contain more males (68.21%) than the general population would suggest (49.20%). Consistent with MTurk, the

sample is also more highly skewed towards lower ages with nearly half of the sample (48.21%) in the 25 to 34 age range, and participants were more highly educated than the general population. Participants reported all levels of income with the highest concentration (27.69%) in the $40,000 to $60,000 range, which is consistent with the "average" level of income used in this study.

[Insert Table 2 Here]

The demographics also indicate that these participants are good proxies for non-sophisticated investors (Buchheit et al. 2018). A significant majority of participants (73.85%) indicated that they own stock in some capacity, indicating that they not only possess the basic knowledge necessary to pass the screening questions but also have some amount about stock ownership experience. Of the participants who indicated they are stock owners, a majority (59.72%) reported owning individual shares, and many owned through a retirement account (44.44%) or an employee stock plan (30.56%). Additionally, many of the participants (37.50%) who reported owning stock also reported having voted as part of a shareholders' meeting either in person or by proxy.

Descriptive Statistics

Descriptive statistics for the observed variables and the scale item indicators are summarized in Table 2. The descriptive statistics are provided for the experimental condition variables from the Dodd-Frank pay ratio as continuous variables. An index of the attractiveness of the investment scale is created by taking the simple average of the five scale items. This index

will only be used for the preliminary analysis as the full measurement model will be used in the SEM analysis of the full model and indirect effects. One item of note is the mean for the propensity to vote against in the say on pay, which indicates that 41% of participants voted against. This is consistent with higher observed levels of dissent for the say on pay, which suggests that the experimental design is increasing the salience of the say on pay vote but not excessively. The scale items indicators consistently demonstrate means near the midpoint of the scale, and they exhibit significant and consistent variance.

[Insert Table 2 Here]

The conditional means for the primary dependent variables of interest are presented as matrix tables in Table 4. The matrix tables cover all nine individual conditions created by the intersection of the PEO compensation and median compensation variables. Means are also tabulated for each level of PEO compensation and median compensation across all levels of the other factor. An addendum to each matrix table tabulates the mean for each ratio condition across all conditions with that ratio.[2]

[Insert Table 3 Here]

Item correlations are provided on Table 3. The correlations for the experimental conditions are included as the continuous variables for brevity and ease of interpretation. All of

[2] The manipulation checks for the manipulated conditions were tested with an ANOVA analysis. The three individual conditions (low, average, high) for both PEO compensation and median compensation were each significantly different. The five factors for the pay ratio were only significantly different in three groups consisting of the highest ratio and high ratio, the low and lowest ratios, and the average ratio.

the scale item indicators are included in the correlation table to confirm their functioning as scale items and to allow for the replication of the structural equation model.

[Insert Table 4 Here]

Structural Equation Model for Opposition to the Say on Pay

The results are further tested with a Structural Equation Modeling (SEM) approach to fully and simultaneously test the fit of the predicted model and the associated indirect effects. The use of SEM also allows for the full inclusion and simultaneous evaluation of the measurement models and the path model, which allows for a more detailed examination of the observed indicators and latent variables. With this approach, all of the hypotheses will be tested including the relevant indirect effects of the mediators on employee sentiments and investor decision making. The models are estimated with MPlus software using maximum likelihood estimation. The indirect effects are tested with bias-corrected bootstrapping to account for the non-linearity of test statistics for the combined path coefficients. The structural equation model is tested with continuous variables for the PEO compensation, median employee compensation, and the pay ratio and with opposition to the say on pay. This model specification will facilitate the estimation and interpretation of the indirect effects.[3]

The analysis of the hypothesized structural equation model is specified with the substitution of the continuous independent variables for the categorical variables of the manipulated conditions. Continuous variables for the independent manipulated factors will

[3] The results of the analysis are robust to other specifications of the model. The results are not significantly different if the pay ratio disclosure variables are specified as categorical conditions with the pay ratio as interactions. They are also not significantly different if the continuous pay ratio disclosure variables are used to predict the propensity to vote no in the say on pay as a categorical outcome (using logistic regression).

simplify the analysis and facilitate the estimation of indirect effects. However, it should first be established that the transformation of the manipulated conditions into the continuous factors does not meaningfully change their interpretation. The preliminary analyses provided limited evidence that the median employee compensation was functioning under this continuous interpretation and mixed, inconclusive evidence regarding the PEO compensation and the pay ratio.

This specification of the structural equation model with the continuous independent variables also demonstrates good fit across all indices. The Chi Square is significant (χ^2 = 383.250, df = 156, p = < 0.001), although this is not unexpected for a model of this size and complexity. The RMSEA (0.086, 90% CI: 0.076 to 0.097) indicates a relatively good global fit with the entire confidence interval below the suggested 0.100. The CFI (0.913) suggests a sufficiently good relative fit, whereas the TLI (0.896) being below 0.900 suggests that the model may marginally achieve sufficient relative fit. Finally, the SRMR (0.062) below 0.800 suggests that the model has good absolute fit (Wang and Wang 2012).

[Insert Table 5 Here]

The results of the measurement model reported in Table 17, Panel B indicate that the measurement of the latent factors for the attractiveness of the investment, fairness, and labor market failures function as expected based on the specification of the scales and that the measurement is not affected by the inclusion of the path model. There were no significant areas of ill fit in the resulting structural equation model, and so no modifications were made to the hypothesized model to achieve model fit. This suggests that the hypothesized path model sufficiently accounts for the significant effects among the variables.

Effects of the Pay Ratio Disclosure on Investor Perceptions and Say on Pay Decisions

The results of the path analysis have been summarized in Figure 4 to facilitate comparison to the hypothesized path model in Figure 3 and the evaluation of the hypotheses. The standardized estimates of the direct effects are drawn from Panels B and C of Table 5. The standardized parameter estimates for the indirect and total effects are not included in Figure 4 but can be found in Panel D of Table 5.

[Insert Figure 4 Here]

Compensation Disclosures as a Signal of Value

If a compensation disclosure held unique, value-relevant information about the productive value of the employee being compensated, there should be a significant direct path from that compensation to an outcome variable representing the investor evaluation or decision, even in the presence of the mediators. Neither H1a (β_{1a} = -0.044, p = 0.362) nor H1b (β_{1b} = 0.066, p = 0.196) are supported, suggesting that investors do not directly relate compensation levels into their assessment of the attractiveness of the investment. Investors similarly failed to impute the level of executive compensation into their decision in the say on pay vote (β_{1c} = 0.017, p = 0.779), which suggests that investors are unable to directly evaluate the productive value of the executive. These results are collectively consistent with the inability of investors to adequately interpret the signal of the economic value of compensation from any associated noise.

Heurist Processing through Labor Market Failures and Fairness

The posited heuristic does explain how investors can process the compensation information from the Dodd-Frank pay ratio disclosure to make decisions in the say on pay and other investment decisions. However, not all of the information is fully incorporated into these factors. Both executive compensation and median compensation are relevant, but there is no incremental effect of the pay ratio.

Perceptions of a labor market failure are dependent on both executive compensation (β_{2a} = 0.354, p = < 0.001) and median compensation (β_{2b} = -0.370, p = < 0.001). However, there is no incremental effect of the pay ratio (β_{2c} = 0.021, p = 0.797). Consistent with theories of market power, investors seem to have perceptions that the labor markets have failed with significantly high levels of executive compensation and/or significantly low levels of median compensation. The lack of an interaction effect represented by the pay ratio suggests that investors process these two separately rather than relying on a single metric of pay disparity, which is consistent with the differential functioning of the labor markets.

Perceptions of fairness are affected by the compensation of the median employee and the prior effect of market failures. H3b regarding the positive relationship between median compensation and fairness is supported (β_{3b} = 0.145, p = 0.091). However, H3a regarding executive compensation (β_{3a} = 0.086, p = 0.324) and H3c regarding the pay ratio (β_{3c} = 0.055, p = 0.632) are not supported. Additionally, there is a significant effect of market failures on perceptions of fairness (β_{3d} = -0.582, p = < 0.001). These results suggest that investors primarily rely on their perceptions of labor market failures to determine whether or not the outcomes are fair but that median compensation has an incremental effect on perceptions of fairness.

Investor Perceptions of the Attractiveness of the Investment

The fourth hypothesis predicts that perceptions of labor market failures and fairness are useful to investors in the evaluation of the attractiveness of the investment. With regard to fairness, H4a predicts a positive effect of fairness on the attractiveness of the investment based on economic and affective perceptions related to the violation of market norms and assessments of the adequacy of compensation. The result of this path in the full path model is significant (β_{4a} = 1.530, p = < 0.001). The effect of labor market failures is significant (β_{4b} = 0.476, p = < 0.001) but not of the expected sign. However, this relationship is impacted by the effect of labor market failures on perceptions of fairness, and so the total effect of labor market failures is negative when considering the indirect effect ($\beta_{3d} \times \beta_{4a}$ = -0.995, 95% CI -1.332 : -0.634)

These results can be extended to evaluate the mediation of the heuristic with the pay ratio disclosure information. Although there are some significant indirect effects of PEO compensation, they are countervailing and do not result in a significant total effect of PEO compensation on the attractiveness of the investments. The direct effect of median compensation on attractiveness was not significant, but its total effect is through the indirect effects of the heuristic. The indirect effects indicate that a one standard deviation increase in the compensation of the median employee (approximately $98,971) is associated with a 0.247 standard deviation increase in perceptions for the attractiveness of the investment. Finally, no direct effect of the pay ratio was posited, but there was also no indirect effect observed through the heuristic of market failures and fairness. These results suggest that investors are able to extrapolate information or revise future expectations regarding the economic value and risk of the company from the compensation of the median employee but not from PEO compensation.

Investor Decision Making in the Say on Pay Vote

In a standard shareholder vote through the proxy statement, shareholders are presented with the matter to be voted on and asked to vote "for", "against", or "abstain" for that proposal. The say on pay vote is the non-binding proposal for the shareholders to approve or oppose executive compensation packages. The say on pay is designed to monitor and deter excessive executive compensation, and so votes against the say on pay are the most salient for analysis. Although H1c predicted that investors should consider the compensation of the PEO as a signal of value of their productivity, there was no evidence that investors consider the level of PEO compensation directly or incrementally when deciding how to vote on the say on pay.

Hypothesis five predicts that investors will use other perceptions of the company to make their decision in voting for or against the say on pay. First, the effects of the heuristic with labor market failures and fairness are considered as H5a and H5b. Fairness is expected to be negatively associated with opposition to the say on pay vote[4], and market failures are expected to be positively associated. The results of the path model indicate that neither perceptions of fairness (β_{5a} = -0.154, p = 0.426) nor perceptions of market failures (β_{5b} = 0.067, p = 0.509) predict the opposition to approving the say on pay proxy vote, and so neither H5a nor H5b are supported. However, H5c regarding the effect of the attractiveness of the investment on the opposition to approve the say on pay is supported (β_{5c} = -0.490, p = 0.004), which indicates that investors are more likely to support the approval of the say on pay vote when they perceive a higher attractiveness of the investment. These results collectively suggest that investors are making economically-motivated decisions despite the affective considerations implied by the assessment

[4] Opposition to the say on pay as a continuous measure is significantly correlated with the decisions to vote against in the say on pay proxy vote. There are no significant differences between this model and the model using the actual vote against as the dependent variable.

of fairness. It provides evidence that investors are using the information from the pay ratio disclosure but implies that investors need the heuristic to be able to adequately process the information for use in decision making.

VI. CONCLUSION

The Dodd-Frank pay ratio disclosure affects investor perceptions of the attractiveness of their investment and consequently their decisions in say on pay proxy votes. However, the evaluation of information processing by investors suggests that they are using perceptions of labor market failures and fairness as a heuristic to evaluate and summarize the compensation disclosure information. Investors use this heuristic to evaluate the attractiveness of the investment including both financial and affective considerations, and the effects of the pay ratio disclosure elements can be observed through this process. Through this process, the Dodd-Frank pay ratio disclosure does fulfil its goal of providing information to investors for say on pay voting.

The Dodd-Frank pay ratio disclosure information is affecting investor decisions regarding the say on pay vote, although these effects are fully mediated through the attractiveness of the investment. This implies that investors are basing their say on pay voting decision primarily on their perception of the attractiveness of the investment, which could be considered a type of pay-for-performance for the executives in question. In contrast, investors are not observed to be making decisions based on the affective considerations surrounding perceptions of fairness and labor market failures. Thus, investors seems to be rationally evaluating executive management on both the financial and non-financial performance of the company reflected in the attractiveness of the investment.

Limitations

This study is designed to simulate the required conditions of the Dodd-Frank pay ratio disclosure. All reporting companies must report the compensation of the PEO, the compensation of an identified median employee, and the ratio between the two, although some other elements are allowed to be included along with these disclosures. The design of pay ratio disclosure in this experiment used only these required elements and minimized the length of the disclosure to ensure readability for the participants, although precise technical language was maintained to reflect the actual language generally used in these disclosures. The final length of the designed pay ratio disclosure was near the minimum length of the observed pay ratio disclosures. Therefore, although the pay ratio disclosure used in the study is simplified, it is not expected to affect the generalizability of the results.

The participants for this study were recruited through the Amazon Mechanical Turk service. Although they were screened to have a basic knowledge of financial reporting, these participants are proxies for non-sophisticated investors. Therefore, these inferences may not be generalizable to more sophisticated investors, particularly those who have specialized knowledge of executive compensation or labor market economics.

The Usefulness of the Dodd-Frank Pay Ratio Disclosure

This study provides evidence that the Dodd-Frank pay ratio disclosure is useful to investors in making assessments about companies and for deciding how to vote on executive compensation. It also demonstrates that these assessments are a complex process through which investors develop perceptions of the company and how it distributes compensation to facilitate

their perceptions of the investment and shareholder decisions. The information from the Dodd-Frank pay ratio disclosure, particularly the compensation of the PEO and the compensation of the median employee, are processed through perceptions of fairness and labor market failures and used by investors in forming perceptions of the attractiveness of their investment and in deciding to vote against the executive compensation in the say on pay shareholder proposal.

The results of this study regarding the usefulness of the Dodd-Frank pay ratio disclosure should be of interest to investors, to companies, and to regulators. Investors and other external stakeholders will be able to use this study to better understand and use the information in the pay ratio disclosure to assess companies based on their compensation practices. Companies should be interested in these results to understand how investors are likely to use the Dodd-Frank pay ratio disclosure and to optimize the informational content of their pay ratio disclosure with the allowable supplemental disclosures. Finally, these results should be of interest to the SEC and other regulators because they provide evidence that information uniquely required by this disclosure is relevant to investors in assessing companies and making decisions regarding shareholder proposals. Regulators and legislators should also be interested in these results for the shaping of new rules and regulations regarding the disclosure of employee compensation. The relevance and useful of this limited disclosure of compensation suggest that requiring additional compensation disclosures would be of interest to investors and relevant to their decision making.

Suggestions for Future Research

This research is a preliminary look at the basic effects of the information for the Dodd-Frank pay ratio disclosure. Although executive compensation has been available and the subject of research, the compensation of any rank-and-file employees is generally not disclosure by

companies and should be of interest for future research. In evaluating the basic effects of the Dodd-Frank pay ratio disclosure, this study used a simplified pay ratio disclosure that did not include any supplementary or superfluous narrative disclosures that could affect how investors perceive and process the compensation information. Additionally, the subject of compensation for rank-and-file employees represented by the median employee and the pay ratio can be perceived as social issues related to income inequality and the distribution of income, and so socially responsible investors may perceive and interpret this information differently than other investors.

Although this study focuses on the Dodd-Frank pay ratio disclosure and it elements, it raises broader questions concerning the usefulness of disclosures of the compensation of employees as a method for investors and other stakeholders to include the value of labor in their decision-making processes. The combination of capital and labor to yield production is one of the foundational principles of microeconomics (Gans et al. 2011), yet accounting has failed to advance the reporting of labor-related financial information. Accounting was born of the need for business enterprises to record and report on investments of capital and the use thereof, and so many of its foundational principles do well in recording and reporting capital but may fail to adequately do so for labor related improvements and expenditures. This is further complicated by the difficulties in quantifying human capital despite its importance to companies (Ballester et al. 2002; Lajili and Zéghal 2005) and by the ownership of human capital by employees that must be rented by companies.

This study expands on the evidence that human capital (Abeysekera 2008; Subbarao and Zeghal 1997; Wyatt and Frick 2010), labor costs (Lajili and Zéghal 2006; Lajili and Zéghal 2005), and related disclosures are relevant and useful to investors and other stakeholders. The

results of this study suggest that compensation disclosures that reveal the distribution of compensation and allow for the analysis of compensation practices are useful and valuable to investors in their decision-making process regarding companies. The Dodd-Frank Act advances the nature of these disclosures by providing for the first required disclosure of rank-and-file employee compensation, which prevents the selective disclosure of favorable information and omission of unfavorable information that was a concern with prior labor-related disclosures. It is hoped that the successful implementation of the Dodd-Frank pay ratio disclosure will open the discussion to increase the required labor-related disclosures, particularly for rank-and-file employees. Investors will benefit by being better able to assess the value and actions of the companies in which they invest, and society may also benefit by being able to assess the compensation practices of companies and the effects that they have on the broader economy.

REFERENCES

Abeysekera, I. 2008. Motivations behind human capital disclosure in annual reports. *Accounting Forum* 32 (1): 16–29.

Adams, J. S. 1965. Inequity in social exchange. In *Advances in experimental social psychology*, 2:267–299. Elsevier.

Agarwal, N. C. 1981. Determinants of executive compensation. *Industrial Relations: A Journal of Economy and Society* 20 (1): 36–45.

Akerlof, G. A., and J. L. Yellen. 1990. The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics* 105 (2): 255–283.

Baker, H. K., and V. Ricciardi. 2014. *Investor behavior: The psychology of financial planning and investing*. John Wiley & Sons.

Balkin, D. B., and L. R. Gomez-Mejia. 1990. Matching compensation and organizational strategies. *Strategic Management Journal* 11 (2): 153–169.

Ballester, M., J. Livnat, and N. Sinha. 2002. Labor Costs and Investments in Human Capital. *Journal of Accounting, Auditing & Finance* 17 (4): 351–373.

Bebchuk, L. A., and J. M. Fried. 2003. Executive compensation as an agency problem. *Journal of economic perspectives* 17 (3): 71–92.

———. 2005. Pay Without Performance: Overview of the Issues. *Journal of Applied Corporate Finance* 17 (4): 8–23.

Bebchuk, L. A., J. M. Fried, and D. I. Walker. 2002. *Managerial power and rent extraction in the design of executive compensation*. National bureau of economic research.

Ben-Amar, W., and D. Zeghal. 2011. Board of directors' independence and executive compensation disclosure transparency: Canadian evidence. *Journal of Applied Accounting Research* 12 (1): 43–60.

Bhaskar, V., A. Manning, and T. To. 2002. Oligopsony and Monopsonistic Competition in Labor Markets. *Journal of Economic Perspectives* 16 (2): 155–174.

Bhaskar, V., and T. To. 2003. Oligopsony and the distribution of wages. *European Economic Review*: 29.

Boeri, T., and J. Van Ours. 2013. *The economics of imperfect labor markets*. Princeton University Press.

Bothma, C. F. C., and G. Roodt. 2013. The validation of the turnover intention scale. *SA Journal of Human Resource Management* 11 (1).

Buchheit, S., M. M. Doxey, T. Pollard, and S. R. Stinson. 2018. A Technical Guide to Using Amazon's Mechanical Turk in Behavioral Accounting Research. *Behavioral Research in Accounting* 30 (1): 111–122.

Bushman, R. M., R. J. Indjejikian, and A. Smith. 1996. CEO compensation: The role of individual performance evaluation. *Journal of Accounting and Economics* 21 (2): 161–193.

Byrd, J. W., M. F. Johnson, and S. L. Porter. 1998. Discretion in Financial Reporting: The Voluntary Disclosure of Compensation Peer Groups in Proxy Statement Performance Graphs. *Contemporary Accounting Research* 15 (1): 25–52.

Castaldo, S., F. Perrini, N. Misani, and A. Tencati. 2009. The Missing Link Between Corporate Social Responsibility and Consumer Trust: The Case of Fair Trade Products. *Journal of Business Ethics* 84 (1): 1–15.

Chen, C. X., and T. Sandino. 2012. Can Wages Buy Honesty? The Relationship Between Relative Wages and Employee Theft: can wages buy honesty? *Journal of Accounting Research* 50 (4): 967–1000.

Connelly, B. L., S. T. Certo, R. D. Ireland, and C. R. Reutzel. 2011. Signaling Theory: A Review and Assessment. *Journal of Management* 37 (1): 39–67.

Cordeiro, J. J., and R. Veliyath. 2003. Beyond pay for performance: A panel study of the determinants of CEO compensation. *American Business Review* 21 (1): 56.

Core, J. E., R. W. Holthausen, and D. F. Larcker. 1999. Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics*: 36.

Cowherd, D. M., and D. I. Levine. 1992. Product Quality and Pay Equity Between Lower-Level Employees and Top Management: An Investigation of Distributive Justice Theory. *Administrative Science Quarterly* 37 (2): 302.

Craven, M. A. 2020. The Usefulness of the Dodd-Frank Pay Ratio Disclosure. Washington State University.

Custódio, C., M. A. Ferreira, and P. Matos. 2013. Generalists versus specialists: Lifetime work experience and chief executive officer pay. *Journal of Financial Economics* 108 (2): 471–492.

Daily, C. M., J. L. Johnson, A. E. Ellstrand, and D. R. Dalton. 1998. Compensation committee composition as a determinant of CEO compensation. *Academy of Management Journal* 41 (2): 209–220.

Demerjian, P., B. Lev, and S. McVay. 2012. Quantifying Managerial Ability: A New Measure and Validity Tests. *Management Science* 58 (7): 1229–1248.

Devers, C. E., A. A. Cannella, G. P. Reilly, and M. E. Yoder. 2007. Executive Compensation: A Multidisciplinary Review of Recent Developments. *Journal of Management* 33 (6): 1016–1072.

Edmans, A., and X. Gabaix. 2009. Is CEO Pay Really Inefficient? A Survey of New Optimal Contracting Theories. *European Financial Management* 15 (3): 486–496.

Ertimur, Y., F. Ferri, and D. Oesch. 2013. Shareholder Votes and Proxy Advisors: Evidence from Say on Pay: Shareholder Votes and Proxy Advisors. *Journal of Accounting Research* 51 (5): 951–996.

Fatemi, A., A. S. Desai, and J. P. Katz. 2003. Wealth creation and managerial pay: MVA and EVA as determinants of executive compensation. *Global Finance Journal* 14 (2): 159–179.

Faulkender, M., D. Kadyrzhanova, N. Prabhala, and L. Senbet. 2010. Executive Compensation: An Overview of Research on Corporate Practices and Proposed Reforms. *Journal of Applied Corporate Finance* 22 (1): 14.

Finkelstein, S., and B. K. Boyd. 1998. How much does the CEO matter? The role of managerial discretion in the setting of CEO compensation. *Academy of Management journal* 41 (2): 179–199.

Finkelstein, S., and D. C. Hambrick. 1989. Chief executive compensation: A study of the intersection of markets and political processes. *Strategic Management Journal* 10 (2): 121–134.

Fisher, J. G., L. A. Maines, S. A. Peffer, and G. B. Sprinkle. 2005. An Experimental Investigation of Employer Discretion in Employee Performance Evaluation and Compensation. *The Accounting Review* 80 (2): 563–583.

Fitz-enz, J. 1997. It's costly to lose good employees. Workforce, 50. *Work Redesign, MA: Addison Welsley*.

Frölich, M., and G. Haile. 2011. Labour markets in developing countries. *Labour Economics* 18: S2–S6.

Gans, J., S. King, and N. G. Mankiw. 2011. *Principles of microeconomics*. Cengage Learning.

Garen, J. E. 1994. Executive compensation and principal-agent theory. *Journal of political economy* 102 (6): 1175–1199.

Gomez-Mejia, L. R., H. Tosi, and T. Hinkin. 1987. Managerial control, performance, and executive compensation. *Academy of Management journal* 30 (1): 51–70.

Gomez-Mejia, L., and R. M. Wiseman. 1997. Reframing Executive Compensation: An Assessment and Outlook. *Journal of Management* 23 (3): 291–374.

Gong, G., L. Y. Li, and J. Y. Shin. 2011. Relative Performance Evaluation and Related Peer Groups in Executive Compensation Contracts. *The Accounting Review* 86 (3): 1007–1043.

Grabke-Rundell, A., and L. R. Gomez-Mejia. 2002. Power as a determinant of executive compensation. *Human Resource Management Review* 12 (1): 3–23.

Graham, J. R., S. Li, and J. Qiu. 2012. Managerial Attributes and Executive Compensation. *Review of Financial Studies* 25 (1): 144–186.

Greenberg, J. E., and J. A. Colquitt. 2005. Handbook of organizational justice.

Hall, B. J., and J. B. Liebman. 1998. Are CEOs really paid like bureaucrats? *The Quarterly Journal of Economics* 113 (3): 653–691.

Hartzell, J. C., and L. T. Starks. 2003. Institutional Investors and Executive Compensation. *The Journal of Finance* 58 (6): 2351–2374.

Heneman, H. G., and D. P. Schwab. 1985. Pay Satisfaction: Its Multidimentional Nature and Measurement. *International Journal of Psychology* 20 (1): 129–141.

Hill, C. W., and P. Phan. 1991. CEO tenure as a determinant of CEO pay. *Academy of Management journal* 34 (3): 707–717.

Ho, J. L. Y., L.-C. Lee, and A. Wu. 2009. How Changes in Compensation Plans Affect Employee Performance, Recruitment, and Retention: An Empirical Study of a Car Dealership. *Contemporary Accounting Research* 26 (1): 167–199.

Hodson, R. 1989. Gender Differences in Job Satisfaction: Why Aren't Women More Dissatisfied? *The Sociological Quarterly* 30 (3): 385–399.

Hyclak, T., G. Johnes, and R. Thornton. 2004. *Fundamentals of Labor Economics*. Houghton Mifflin.

Indjejikian, R. J. 1999. Performance Evaluation and Compensation Research: An Agency Perspective. *Accounting Horizons* 13 (2): 147–157.

Jensen, M. C., and W. H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of financial economics* 3 (4): 305–360.

Kelly, K., B. M. Low, H.-T. Tan, and S.-K. Tan. 2012. Investors' Reliance on Analysts' Stock Recommendations and Mitigating Mechanisms for Potential Overreliance. *Contemporary Accounting Research* 29 (3).

Kelly, K., and J. L. Seow. 2016. Investor Reactions to Company Disclosure of High CEO Pay and High CEO-to-Employee Pay Ratio: An Experimental Investigation. *Journal of Management Accounting Research* 28 (1): 107–125.

Kiatpongsan, S., and M. I. Norton. 2014. How Much (More) Should CEOs Make? A Universal Desire for More Equal Pay. *Perspectives on Psychological Science* 9 (6): 587–593.

Kimbro, M. B., and D. Xu. 2016. Shareholders have a say in executive compensation: Evidence from say-on-pay in the United States. *Journal of Accounting and Public Policy* 35 (1): 19–42.

Kostiuk, P. F. 1990. Firm Size and Executive Compensation. *The Journal of Human Resources* 25 (1): 90.

Lajili, K., and D. Zéghal. 2005. Labor cost voluntary disclosures and firm equity values: Is human capital information value-relevant? *Journal of International Accounting, Auditing and Taxation* 14 (2): 121–138.

———. 2006. Market performance impacts of human capital disclosures. *Journal of Accounting and Public Policy* 25 (2): 171–194.

Levine, D. I. 1993. Fairness, markets, and ability to pay: Evidence from compensation executives. *The American Economic Review*: 1241–1259.

Main, B. G., C. A. O'Reilly, and J. Wade. 1995. The CEO, the board of directors and executive compensation: Economic and psychological perspectives. *Industrial and Corporate Change* 4 (2): 293–332.

Mallette, P., R. D. Middlemist, and W. E. Hopkins. 1995. Social, political, and economic determinants of chief executive compensation. *Journal of Managerial Issues*: 253–276.

Martin, R. M. 2015. The Effect of Executive Compensation Discretion on Investor Judgments. The University of Wisconsin-Madison.

Mason, W., and S. Suri. 2012. Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods* 44 (1): 1–23.

Mohan, B., M. I. Norton, and R. Deshpande. 2015. Paying Up for Fair Pay: Consumers Prefer Firms with Lower CEO-to-Worker Pay Ratios. *SSRN Electronic Journal*.

O'Reilly, C. A., B. G. Main, and G. S. Crystal. 1988. CEO Compensation as Tournament and Social Comparison: A Tale of Two Theories. *Administrative Science Quarterly* 33 (2): 257.

Phelan, J. 1994. The Paradox of the Contented Female Worker: An Assessment of Alternative Explanations. *Social Psychology Quarterly* 57 (2): 95.

Price, J. L. 1997. Handbook of organizational measurement. *International journal of manpower* 18 (4–1): 305–558.

Ramlall, S. 2004. A Review of Employee Motivation Theories and their Implications for Employee Retention within Organizations. *Journal of the American Academy of Business*: 13.

Redling, E. T. 1981. Myth vs. Reality: The Relationship Between Top Executive Pay and Corporate Performance. *Compensation Review* 13 (4): 16–24.

Scarpello, V., V. Huber, and R. J. Vandenberg. 1988. Compensation satisfaction: Its measurement and dimensionality. *Journal of Applied Psychology* 73 (2): 163.

Securities and Exchange Commission. 2015. Pay Ratio Disclosure; Final Rule. *Federal Register*: 50103–50187.

Sridharan, U. V. 1996. CEO Influence and Executive Compensation. *The Financial Review* 31 (1): 51–66.

Subbarao, A. V., and D. Zeghal. 1997. Human Resources Information Disclosure in Annual Reports: An International Comparison. *Journal of Human Resource Costing & Accounting* 2 (2): 53–73.

Summers, T. P., and W. H. Hendrix. 1991. Modelling the role of pay equity perceptions: A field study. *Journal of Occupational Psychology* 64 (2): 145–157.

Terpstra, D. E., and A. L. Honoree. 2003. The Relative Importance of External,Internal, Individual and Procedural Equity to Pay Satisfaction: Procedural Equity May be More Important to Employees Than Organizations Believe. *Compensation & Benefits Review* 35 (6): 67–74.

Tosi, H. L., S. Werner, J. P. Katz, and L. R. Gomez-Mejia. 2000. How Much Does Performance Matter? A Meta-Analysis of CEO Pay Studies. *Journal of Management* 26 (2): 39.

Vafeas, N. 2003. Further evidence on compensation committee composition as a determinant of CEO compensation. *Financial Management*: 53–70.

Vafeas, N., and Z. Afxentiou. 1998. The association between the SEC's 1992 compensation disclosure rule and executive compensation policy changes. *Journal of Accounting and Public Policy* 17 (1): 27–54.

Wadongo, B., and M. Abdel-Kader. 2014. Contingency theory, performance management and organisational effectiveness in the third sector: A theoretical framework. Ed. D. Luisa D. Huaccho Huatuco, Dr Claire. *International Journal of Productivity and Performance Management* 63 (6): 680–703.

Wang, J., and X. Wang. 2012. Structural Equation Modeling: Applications Using Mplus.

Werner, S., and S. G. Ward. 2004. Recent compensation research: An eclectic review. *Human Resource Management Review* 14 (2): 201–227.

Widener, S. K. 2006. Human capital, pay structure, and the use of performance measures in bonus compensation. *Management Accounting Research* 17 (2): 198–221.

Wyatt, A., and H. Frick. 2010. Accounting for Investments in Human Capital: A Review. *Australian Accounting Review* 20 (3): 199–220.

Yermack, D. 2006. Flights of fancy: Corporate jets, CEO perquisites, and inferior shareholder returns. *Journal of Financial Economics*: 32.

FIGURE 1

Interpretation of Disclosed Compensation in Perfectly Efficient Labor Markets

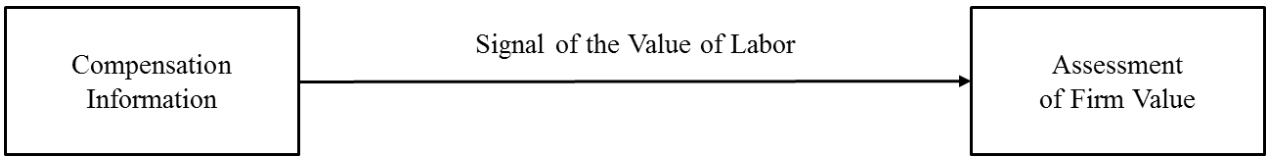
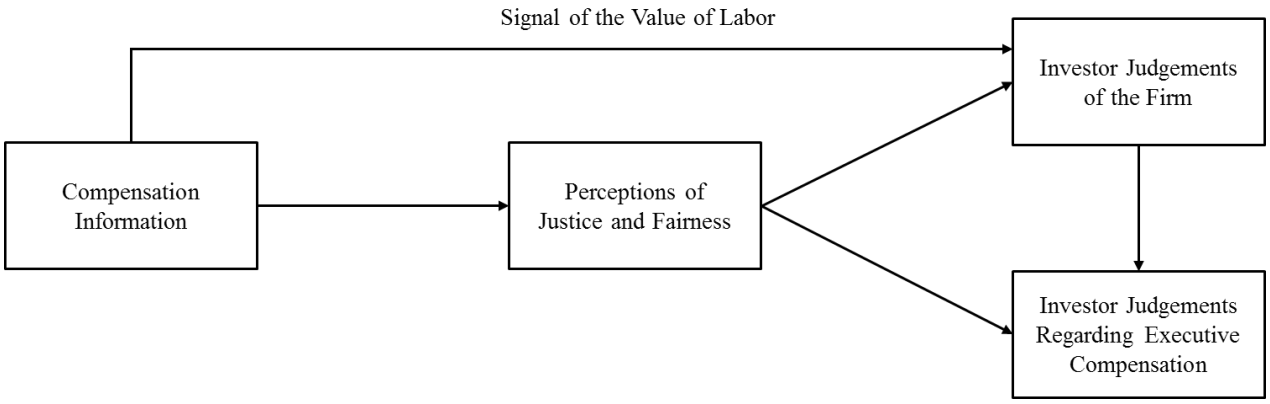


FIGURE 2

Interpretation of Disclosed Compensation with Labor Market Inefficiencies



Page 54

FIGURE 3

Hypothetical Model of Investor Interpretation of the Pay Ratio Disclosure

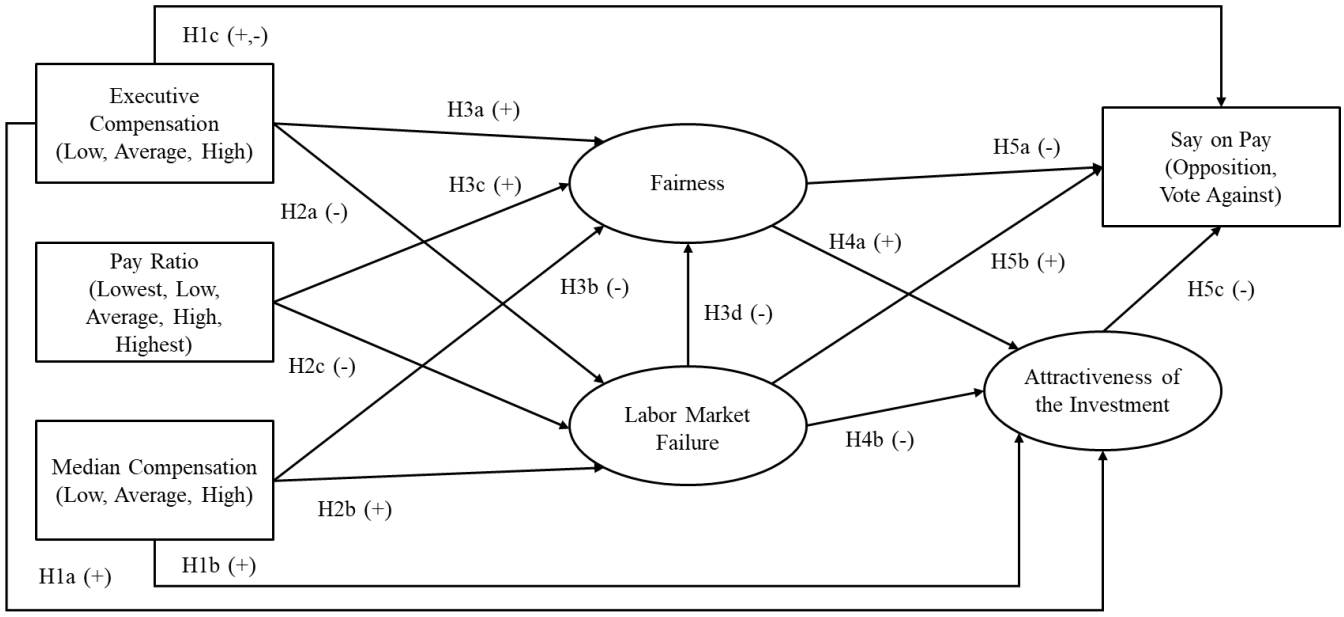
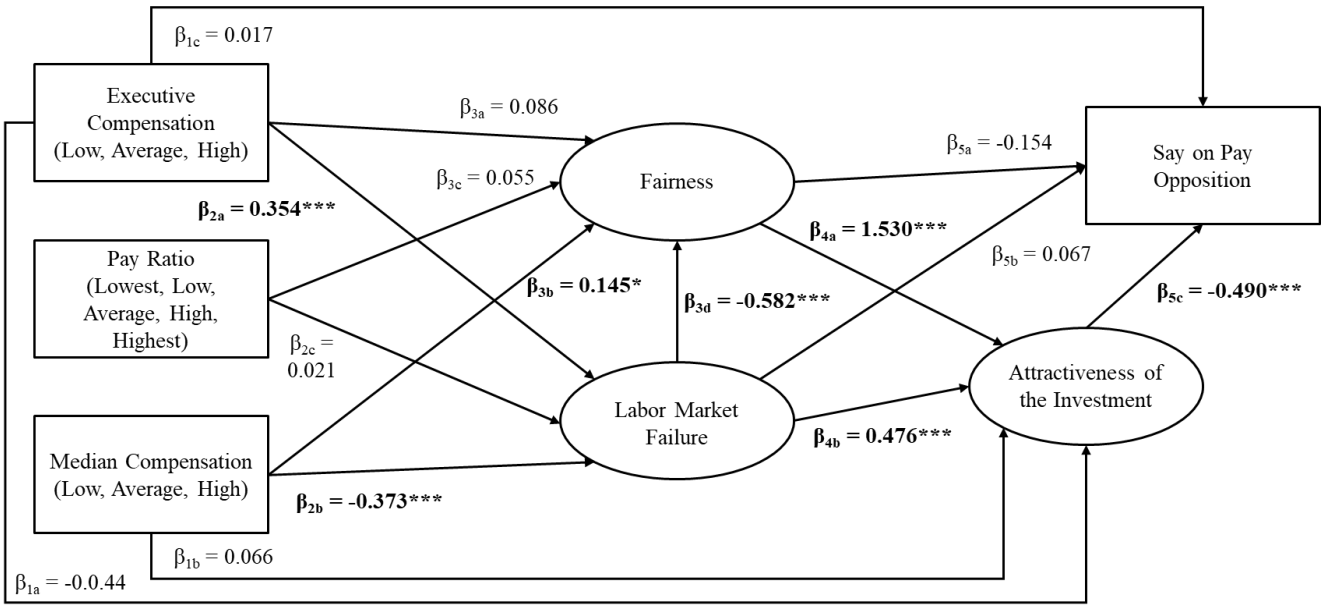


FIGURE 4

Path Results for the Structural Equation Model



*, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.

TABLE 1

Participant Demographics

	Participants (*n=195*)	Percentage	Population (U.S.A 2010)
Gender			
Male	133	68.21%	49.20%
Female	59	30.26%	50.80%
Other	1	0.51%	
Prefer not to say	2	1.03%	
Age			
18 to 24	11	5.64%	6.99%
25 to 34	94	48.21%	13.30%
35 to 44	44	22.56%	13.30%
45 to 54	28	14.36%	14.58%
55 to 64	12	6.15%	11.82%
65 to 74	4	2.05%	7.03%
75 to 84	0	0.00%	4.23%
85 or older	0	0.00%	1.78%
Prefer not to say	2	1.03%	
Educational Attainment			
Some high school	1	0.51%	
High school graduate	16	8.21%	29.50%
Some college	17	8.72%	16.60%
2 year degree	22	11.28%	9.80%
4 year degree	93	47.69%	20.50%
Professional degree	42	21.54%	12.00%
Doctorate	2	1.03%	
Prefer not to say	2	1.03%	

TABLE 1 (Continued)

Participant Demographics

	Participants (*n=195*)	Percentage
Employment Status		
Employed full time	155	79.49%
Employed part time	19	9.74%
Unemployed, looking	8	4.10%
Unemployed, not looking	4	2.05%
Retired	2	1.03%
Student	2	1.03%
Disabled	0	0.00%
Prefer not to say	5	2.56%
Income		
Less than $20,000	23	11.79%
$20,000 to $40,000	46	23.59%
$40,000 to $60,000	54	27.69%
$60,000 to $80,000	35	17.95%
$80,000 to $100,000	16	8.21%
More than $100,000	14	7.18%
Prefer not to say	7	3.59%
Stock Ownership		
Yes	144	73.85%
No	49	25.13%
I don't know	2	1.03%

TABLE 1 (Continued)

Participant Demographics

	Participants (*n=144*)	Percentage
Ownership Type (*Independent*)		
Individually	86	59.72%
Employee plan	44	30.56%
Managed fund	49	34.03%
Retirement account	64	44.44%
Other	2	1.39%
Prefer not to say	3	2.08%
Proxy Voting Experience		
Yes	54	37.50%
No	89	61.81%
I don't know	1	0.69%

Population demographics are taken from the 2010 United States Census data

TABLE 2

Descriptive Statistics

Panel A: Descriptive Statistics of the Primary Variables

Variable	n	Mean	Std. Dev.	Q1	Median	Q3
PEO Comp.	195	14.697	15.009	1.510	7.382	36.083
Median Comp.	195	97.642	98.971	9.856	48.176	235.485
Pay Ratio	195	0.619	1.095	0.031	0.153	0.749

Variable	n	Mean	Std. Dev.	Q1	Median	Q3
Attract Index	195	4.37	1.44	3.4	4.6	5.4
SoP Support	195	4.30	2.00	2.0	5.0	6.0
SoP Opposition	195	3.70	2.00	6.0	5.0	2.0
SoP Vote Against	195	0.41	0.49	0.0	0.0	1.0

Panel B: Descriptive Statistics of the Measurement Model Indicators

Variable	n	Mean	Std. Dev.	Q1	Median	Q3
Attractive 1	195	4.55	1.59	4.0	5.0	6.0
Attractive 2	195	4.24	1.76	3.0	4.0	5.0
Attractive 3	195	4.18	1.62	3.0	4.0	5.0
Attractive 4	195	4.74	1.62	4.0	5.0	6.0
Attractive 5	195	4.16	1.60	3.0	4.0	5.0
Fairness 1	195	4.39	1.94	3.0	5.0	6.0
Fairness 2	195	4.42	1.71	3.0	5.0	6.0
Fairness 3	195	4.16	2.01	2.0	5.0	6.0
Fairness 4	195	4.18	1.79	3.0	4.0	6.0
Fairness 5	195	4.03	1.99	2.0	4.0	6.0
Fairness 6	195	3.79	1.82	2.0	4.0	5.0
Fairness 7	195	3.88	1.98	2.0	4.0	6.0
Market Failure 1	195	3.34	1.78	2.0	3.0	5.0
Market Failure 2	195	3.37	1.95	2.0	3.0	5.0
Market Failure 3	195	3.49	2.01	2.0	3.0	5.0
Market Failure 4	195	2.86	1.88	1.0	2.0	4.0

TABLE 3

Correlation Matrix

		(1)	(2)	(3)	(4)	(5)	(6)
(1)	*PEO Comp*	*1.000*					
(2)	*Median Comp*	0.005	*1.000*				
(3)	*Pay Ratio*	**0.578**	**-0.425**	*1.000*			
(4)	*Attract Index*	-0.037	**0.289**	-0.123	*1.000*		
(5)	*SoP Opposition*	-0.071	0.137	-0.109	**0.619**	*1.000*	
(6)	*SoP Against*	0.039	-0.111	0.052	**-0.470**	**0.729**	*1.000*
(7)	*Attractiveness 1*	-0.122	**0.263**	**-0.167**	**0.896**	**-0.529**	**-0.397**
(8)	*Attractiveness 2*	0.042	**0.273**	-0.056	**0.918**	**-0.582**	**-0.410**
(9)	*Attractiveness 3*	-0.020	**0.256**	-0.110	**0.892**	**-0.553**	**-0.405**
(10)	*Attractiveness 4*	-0.094	**0.206**	-0.105	**0.861**	**-0.596**	**-0.488**
(11)	*Attractiveness 5*	0.025	**0.265**	-0.110	**0.812**	**-0.449**	**-0.358**
(12)	*Fairness 1*	**-0.216**	**0.208**	**-0.205**	**0.667**	**-0.516**	**-0.382**
(13)	*Fairness 2*	**-0.192**	**0.141**	-0.135	**0.647**	**-0.557**	**-0.437**
(14)	*Fairness 3*	-0.063	**0.353**	**-0.188**	**0.648**	**-0.446**	**-0.363**
(15)	*Fairness 4*	-0.063	**0.254**	**-0.167**	**0.674**	**-0.517**	**-0.386**
(16)	*Fairness 5*	-0.013	**0.376**	-0.133	**0.729**	**-0.545**	**-0.415**
(17)	*Fairness 6*	-0.063	**0.143**	-0.079	**0.648**	**-0.535**	**-0.339**
(18)	*Fairness 7*	-0.087	**0.299**	**-0.149**	**0.695**	**-0.491**	**-0.379**
(19)	*Market Failure 1*	**-0.288**	**0.165**	**-0.237**	**0.157**	**-0.195**	**-0.204**
(20)	*Market Failure 2*	**-0.451**	**0.200**	**-0.363**	**0.193**	**-0.227**	**-0.247**
(21)	*Market Failure 3*	**-0.170**	**0.391**	**-0.315**	**0.162**	-0.107	-0.138
(22)	*Market Failure 4*	**-0.222**	**0.317**	**-0.245**	**0.348**	**-0.355**	**-0.294**

TABLE 3 (Continued)

Correlation Matrix

		(7)	(8)	(9)	(10)	(11)	(12)
(7)	*Attractiveness 1*	*1.000*					
(8)	*Attractiveness 2*	**0.784**	*1.000*				
(9)	*Attractiveness 3*	**0.749**	**0.778**	*1.000*			
(10)	*Attractiveness 4*	**0.756**	**0.729**	**0.730**	*1.000*		
(11)	*Attractiveness 5*	**0.639**	**0.713**	**0.650**	**0.557**	*1.000*	
(12)	*Fairness 1*	**0.583**	**0.621**	**0.591**	**0.608**	**0.517**	*1.000*
(13)	*Fairness 2*	**0.631**	**0.570**	**0.543**	**0.619**	**0.473**	**0.595**
(14)	*Fairness 3*	**0.533**	**0.598**	**0.546**	**0.582**	**0.578**	**0.661**
(15)	*Fairness 4*	**0.593**	**0.626**	**0.589**	**0.629**	**0.514**	**0.577**
(16)	*Fairness 5*	**0.611**	**0.725**	**0.613**	**0.640**	**0.598**	**0.667**
(17)	*Fairness 6*	**0.583**	**0.612**	**0.572**	**0.544**	**0.523**	**0.549**
(18)	*Fairness 7*	**0.605**	**0.667**	**0.592**	**0.609**	**0.568**	**0.687**
(19)	*Market Failure 1*	**0.148**	0.087	0.121	**0.250**	0.086	**0.151**
(20)	*Market Failure 2*	**0.186**	0.077	**0.177**	**0.300**	0.113	**0.393**
(21)	*Market Failure 3*	**0.159**	0.123	0.117	**0.214**	0.096	**0.308**
(22)	*Market Failure 4*	**0.299**	**0.305**	**0.272**	**0.417**	**0.231**	**0.424**

TABLE 3 (Continued)

Correlation Matrix

		(13)	(14)	(15)	(16)	(17)	(18)
(13)	*Fairness 2*	*1.000*					
(14)	*Fairness 3*	**0.519**	*1.000*				
(15)	*Fairness 4*	**0.611**	**0.715**	*1.000*			
(16)	*Fairness 5*	**0.580**	**0.802**	**0.696**	*1.000*		
(17)	*Fairness 6*	**0.552**	**0.588**	**0.686**	**0.660**	*1.000*	
(18)	*Fairness 7*	**0.615**	**0.789**	**0.742**	**0.825**	**0.664**	*1.000*
(19)	*Market Failure 1*	**0.171**	**0.207**	**0.179**	**0.229**	**0.144**	**0.186**
(20)	*Market Failure 2*	**0.270**	**0.298**	**0.194**	**0.283**	0.139	**0.225**
(21)	*Market Failure 3*	**0.142**	**0.489**	**0.309**	**0.436**	**0.238**	**0.372**
(22)	*Market Failure 4*	**0.387**	**0.554**	**0.428**	**0.570**	**0.367**	**0.520**

		(19)	(20)	(21)	(22)
(19)	*Market Failure 1*	*1.000*			
(20)	*Market Failure 2*	**0.583**	*1.000*		
(21)	*Market Failure 3*	**0.403**	**0.494**	*1.000*	
(22)	*Market Failure 4*	**0.492**	**0.582**	**0.596**	*1.000*

Correlations indicated in bold are significant at the 5% level.

TABLE 4

Conditional Means of the Opposition to the Say on Pay

		Below-Average PEO Compensation $1,510,215	*Average PEO Compensation $7,381,932*	*Above-Average PEO Compensation $36,082,882*
	All Conditions **3.70** (1.44) *n=195*	*Below-Average PEO Total* **3.37** (1.92) *n=65*	*Average PEO Total* **3.87** (1.98) *n=67*	*Above-Average PEO Total* **3.86** (1.08) *n=63*
Below-Average Median Compensation $9,856	*Below-Average Median Total* **4.27** (2.05) *n=66*	*Average Ratio 153 to 1* **4.26** (1.81) *n=23*	*High Ratio 749 to 1* **4.50** (2.15) *n=22*	*Highest Ratio 3,661 to 1* **4.05** (2.25) *n=21*
Average Median Compensation $48,176	*Average Median Total* **3.42** (1.97) *n=64*	*Low Ratio 31 to 1* **2.38** (1.47) *n=21*	*Average Ratio 153 to 1* **3.82** (1.89) *n=22*	*High Ratio 749 to 1* **4.05** (2.16) *n=21*
Above-Average Median Compensation $235,485	*Above-Average Median Total* **3.38** (1.86) *n=65*	*Lowest Ratio 6 to 1* **3.38** (2.01) *n=21*	*Low Ratio 31 to 1* **3.30** (1.79) *n=23*	*Average Ratio 153 to 1* **3.48** (1.86) *n=*

Lowest Ratio Total **3.38** (2.01) *n=21*	*Low Ratio Total* **2.86** (1.69) *n=44*	*Average Ratio Total* **3.86** (1.86) *n=66*	*High Ratio Total* **4.28** (2.14) *n=43*	*Highest Ratio Total* **4.05** (2.25) *n=21*

TABLE 5

Structural Equation Modeling (SEM) Results for Opposition to Say on Pay

Panel A: Model Fit Indices

Chi Square			RMSEA					
Value	**df**	**p-value**	**Estimate**	**90% CI**		**CFI**	**TLI**	**SRMR**
383.250	156	< 0.001	0.086	0.076	0.097	0.913	0.896	0.062

TABLE 5

Structural Equation Modeling (SEM) Results for Opposition to Say on Pay

Panel B: Standardized Parameter Estimates for the Measurement Model

Path			Estimate	Std. Error	p-value
Attractiveness	by	Attractiveness 1	**0.864**	**0.024**	**< 0.001**
Attractiveness	by	Attractiveness 2	**0.912**	**0.019**	**< 0.001**
Attractiveness	by	Attractiveness 3	**0.858**	**0.027**	**< 0.001**
Attractiveness	by	Attractiveness 4	**0.826**	**0.032**	**< 0.001**
Attractiveness	by	Attractiveness 5	**0.754**	**0.036**	**< 0.001**
Fairness	by	Fairness 1	**0.764**	**0.042**	**< 0.001**
Fairness	by	Fairness 2	**0.693**	**0.049**	**< 0.001**
Fairness	by	Fairness 3	**0.858**	**0.024**	**< 0.001**
Fairness	by	Fairness 4	**0.814**	**0.029**	**< 0.001**
Fairness	by	Fairness 5	**0.904**	**0.020**	**< 0.001**
Fairness	by	Fairness 6	**0.744**	**0.038**	**< 0.001**
Fairness	by	Fairness 7	**0.898**	**0.021**	**< 0.001**
Market Failure	by	Market Failure 1	**0.608**	**0.078**	**< 0.001**
Market Failure	by	Market Failure 2	**0.725**	**0.066**	**< 0.001**
Market Failure	by	Market Failure 3	**0.715**	**0.055**	**< 0.001**
Market Failure	by	Market Failure 4	**0.832**	**0.044**	**< 0.001**

TABLE 5 (Continued)

Structural Equation Modeling (SEM) Results for Opposition to Say on Pay

Panel C: Standardized Parameter Estimates for the Path Model

Path			Estimate		Std. Error	p-value
Market Failure	on	PEO Comp	**0.354**	***	**0.082**	**< 0.001**
Market Failure	on	Median Comp	**-0.370**	***	**0.088**	**< 0.001**
Market Failure	on	Pay Ratio	0.021		0.081	0.797
Fairness	on	PEO Comp	0.086		0.087	0.324
Fairness	on	Median Comp	**0.145**	*	**0.086**	**0.091**
Fairness	on	Pay Ratio	0.055		0.114	0.632
Fairness	on	Market Failure	**-0.582**	***	**0.089**	**< 0.001**
Attractiveness	on	PEO Comp	-0.044		0.049	0.362
Attractiveness	on	Median Comp	0.066		0.051	0.196
Attractiveness	on	Market Failure	**0.283**	***	**0.076**	**< 0.001**
Attractiveness	on	Fairness	**0.975**	***	**0.054**	**< 0.001**
SoP Opposition	on	PEO Comp	0.017		0.060	0.779
SoP Opposition	on	Market Failure	0.067		0.101	0.509
SoP Opposition	on	Fairness	-0.154		0.194	0.426
SoP Opposition	on	Attractiveness	**-0.490**	***	**0.172**	**0.004**

TABLE 5 (Continued)

Structural Equation Modeling (SEM) Results for Opposition to Say on Pay

Panel D: Mediation Tests

Structural Relationships	Estimate	Std. Error	p-value	Bootstrap 95% CI	
Effects of PEO Comp on Attractiveness					
Direct Effect of PEO Compensation	-0.044	0.049	0.362	-0.130	0.032
Indirect Effects					
Through Fairness	0.084	0.087	0.331	-0.057	0.230
Through Market Failure	**0.100**	**0.036**	**0.006**	**0.052**	**0.174**
Through Market Failure and Fairness	**-0.201**	**0.062**	**0.001**	**-0.319**	**-0.115**
Sum of Indirect Effects	-0.017	0.080	0.836	-0.149	0.117
Total Effect of PEO Comp on Attractiveness	-0.061	0.086	0.478	-0.201	0.082

Structural Relationships	Estimate	Std. Error	p-value	Bootstrap 95% CI	
Effects of Median Comp on Attractiveness					
Direct Effect of PEO Compensation	0.066	0.051	0.196	-0.013	0.154
Indirect Effects					
Through Fairness	0.141	0.083	0.087	-0.001	0.270
Through Market Failure	**-0.105**	**0.042**	**0.013**	**-0.189**	**-0.050**
Through Market Failure and Fairness	**0.210**	**0.073**	**0.004**	**0.106**	**0.344**
Sum of Indirect Effects	**0.247**	**0.069**	**< 0.001**	**0.134**	**0.360**
Total Effect of Median Comp on Attractiveness	**0.313**	**0.074**	**< 0.001**	**0.188**	**0.430**

TABLE 5 (Continued)

Structural Equation Modeling (SEM) Results for Opposition to Say on Pay

Panel D: Mediation Tests (Continued)

Structural Relationships	Estimate	Std. Error	p-value	Bootstrap 95% CI	
Effects of Pay Ratio on Attractiveness					
Direct Effect of PEO Compensation					
Indirect Effects					
Through Fairness	0.054	0.111	0.631	-0.113	0.254
Through Market Failure	0.006	0.024	0.805	-0.033	0.045
Through Market Failure and Fairness	-0.012	0.047	0.802	-0.089	0.066
Sum of Indirect Effects	0.048	0.113	0.673	-0.124	0.246
Total Effect of Pay Ratio on Attractiveness	0.048	0.113	0.673	-0.124	0.246

TABLE 5 (Continued)

Structural Equation Modeling (SEM) Results for Opposition to Say on Pay

Panel D: Mediation Tests

Structural Relationships	Estimate	Std. Error	p-value	Bootstrap 95% CI	
Effects of PEO Comp on SoP Opposition					
Direct Effect of PEO Compensation	0.017	0.060	0.779	-0.080	0.114
Indirect Effects					
Through Attractiveness	0.022	0.027	0.413	-0.001	0.011
Through Fairness	-0.013	0.028	0.629	-0.004	0.012
Through Fairness and Attractiveness	-0.041	0.047	0.380	-0.019	0.002
Through Market Failure	0.024	0.037	0.527	-0.032	0.092
Through Market Failure and Attractiveness	**-0.049**	**0.025**	**0.053**	**-0.106**	**-0.019**
Through Market Failure and Fairness	0.032	0.043	0.458	-0.024	0.117
Through Market Failure, Fairness, and Attractiveness	**0.098**	**0.046**	**0.032**	**0.040**	**0.195**
Sum of Indirect Effects	0.050	0.061	0.413	-0.006	0.020
Total Effect of PEO Comp on SoP Opposition	0.089	0.077	0.247	-0.005	0.028

TABLE 5 (Continued)

Structural Equation Modeling (SEM) Results for Opposition to Say on Pay

Panel D: Mediation Tests (Continued)

Structural Relationships	Estimate	Std. Error	p-value	Bootstrap 95% CI	
Effects of Median Comp on SoP Opposition					
Direct Effect of Median Comp					
Indirect Effects					
Through Attractiveness	-0.032	0.026	0.210	-0.086	0.000
Through Fairness	-0.022	0.034	0.508	-0.110	0.010
Through Fairness and Attractiveness	**-0.069**	**0.052**	**0.180**	**-0.178**	**-0.004**
Through Market Failure	-0.025	0.038	0.519	-0.091	0.035
Through Market Failure and Attractiveness	**0.051**	**0.028**	**0.071**	**0.019**	**0.114**
Through Market Failure and Fairness	-0.033	0.046	0.470	-0.125	0.024
Through Market Failure, Fairness, and Attractiveness	**-0.103**	**0.051**	**0.046**	**-0.207**	**-0.038**
Sum of Indirect Effects	**-0.234**	**0.056**	**< 0.001**	**-0.326**	**-0.147**
Total Effect of Median Comp on SoP Opposition	**-0.234**	**0.056**	**< 0.001**	**-0.326**	**-0.147**

TABLE 5 (Continued)

Structural Equation Modeling (SEM) Results for Opposition to Say on Pay

Panel D: Mediation Tests (Continued)

Structural Relationships	Estimate	Std. Error	p-value	Bootstrap 95% CI	
Effects of Pay Ratio on SoP Opposition					
Direct Effect of Pay Ratio					
Indirect Effects					
Through Fairness	-0.008	0.029	0.773	-0.104	0.012
Through Fairness and Attractiveness	-0.026	0.060	0.660	-0.157	0.047
Through Market Failure	0.001	0.010	0.890	-0.007	0.028
Through Market Failure and Attractiveness	-0.003	0.013	0.818	-0.026	0.014
Through Market Failure and Fairness	0.002	0.012	0.879	-0.008	0.033
Through Market Failure, Fairness, and Attractiveness	0.006	0.025	0.813	-0.030	0.050
Sum of Indirect Effects	-0.029	0.076	0.707	-0.159	0.088
Total Effect of Pay Ratio on SoP Opposition	-0.029	0.076	0.707	-0.159	0.088

For the measurement and path model parameter estimates in Panels B and C, *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively. Indirect effects in Panel D are tested at the 5% level of significance with a bias-corrected bootstrapped confidence interval based on a bootstrapped sample of 5000. “By” paths are the effects of the latent factors on the reflective indicators used to measure those latent variables. “On” paths represent regression effects of the second variable on the first.

APPENDIX A

Variable Definitions

Variables	Definition
PEO Comp	Compensation of the PEO in millions of dollars.
Median Comp	Compensation of the median employee in thousands of dollars.
Pay Ratio	Pay ratio of the PEO to the median employee in thousandths. (Equal to the PEO Compensation variable divided by the Median Compensation variable.)
Attractiveness Index	Simple average of the Attractiveness of the investment items. Used in the preliminary ANOVA analyses.
SoP Support	Support for approving the say on pay vote regarding executive compensation. Measured on a seven-point Likert-type scale from "Strongly Oppose" to "Strongly Support".
SoP Opposition	Opposition to approving the say on pay vote regarding executive compensation. Reverse coding of Say on Pay Support.
SoP Vote Against	Voting response to the say on pay vote. Dichotomous variable set to 1 if "Against" is selected and set to 0 if "For" or "Abstain" are selected.
Attractiveness 1	First item of the Attractiveness of the Investment scale. "Melba Marketing is an attractive investment." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Attractiveness 2	Second item of the Attractiveness of the Investment scale. "I would consider increasing my investment in Melba Marketing." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Attractiveness 3	Third item of the Attractiveness of the Investment scale. "I would recommend Melba Marketing as an investment to my friends and family." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Attractiveness 4	Fourth item of the Attractiveness of the Investment scale. "I am planning to maintain my investment in Melba Marketing." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Attractiveness 5	Fifth item of the Attractiveness of the Investment scale. "Melba Marketing is more valuable than its financial performance would suggest." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".

APPENDIX A (Continued)

Variable Definitions

Variables	Definition
Fairness 1	First item of the Fairness scale. "Compared to other CEOs, the compensation of Melba Marketing's CEO is fair." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Fairness 2	Second item of the Fairness scale. "Melba Marketing's executive compensation is consistent with the financial performance of the company." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Fairness 3	Third item of the Fairness scale. "Compared to other rank-and-file employees, the compensation of Melba Marketing's median employee is fair." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Fairness 4	Fourth item of the Fairness scale. "The compensation of Melba Marketing's rank-and-file employees is consistent with their productivity." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Fairness 5	Fifth item of the Fairness scale. "Melba Marketing compensates its employees fairly across all levels of the organization." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Fairness 6	Sixth item of the Fairness scale. "The difference in pay between the CEO of Melba Marketing and its median employee reflects their contributions to the company." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Fairness 7	Seventh item of the Fairness scale. "The compensation policies of Melba Marketing are equitable to both its CEO and rank-and-file employees." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Market Failure 1	First item of the Market Failure scale. "The CEO of Melba Marketing used their power to receive above-market compensation." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Market Failure 2	Second item of the Market Failure scale. "Compared to other CEOs, the CEO of Melba Marketing is overpaid." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".

APPENDIX A (Continued)

Variable Definitions

Variables	Definition
Market Failure 3	Third item of the Market Failure scale. "Compared to other median employees, the median employee of Melba Marketing is underpaid." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".
Market Failure 4	Fourth item of the Market Failure scale. "The pay ratio between Melba Marketing's CEO and its median employee is excessive." Measured on a seven-point Likert-type scale from "Strongly disagree" to "Strongly agree".

Linking Efficiency, Profitability and Productivity Measures, with Application to the U.S. Airline Industry

ABSTRACT

Improved operational efficiency is widely recognized as enhancing organizational competitiveness, productivity, and profitability. Yet a direct linkage among measures of these concepts is hardly evident in fields that logically or theoretically address these concepts, including finance, economics, and operations management.

This paper suggests a way in which a specific connection can be established among measures of these concepts using a multidisciplinary setting in which appropriate measures from different but related fields are linked. These linkages are then explored in an empirical setting using time series-cross-sectional data from the U.S. Airline industry data.

Introduction

Improved operational efficiency is widely recognized as enhancing organizational competitiveness, productivity, and profitability. Yet a direct linkage among measures of these concepts is hardly evident in fields that logically or theoretically address these concepts, including finance, economics, and operations management. Efficiency and productivity, moreover, have been said to be different concepts, implying that gains in one need not create similar movement in the other. Empirical evidence, too, can be inconsistent. Some companies or industries, for example, have experienced impressive productivity gains without a commensurate gain in profitability (e.g. airlines). National economic data, too, at times indicates a divergence between productivity growth and corporate earnings (see Appendix Table 1).

This paper suggests a way in which a specific connection can be established among measures of these concepts using a multidisciplinary setting in which appropriate measures from different but related fields are linked. To this end, it is shown that a common economic measure of improved efficiency, lower average (unit) cost, can be linked to financial measures of the rate and level of profitability. These measures can also be linked to multifactor productivity, using a specification of the latter from a well-known text in Operations Management. These linkages are

then explored in an empirical setting using time series-cross-sectional data from the U.S. Airline industry data.

In this section, we consider the linkage between a measure of financial efficiency and two specific measures of profitability, the profitability percentage and profitability margin. Then, we address the mathematical connection between this efficiency measure and both the break-even and the target-profit level of output, in the framework of the basic cost-volume-profit (CVP) model.

A well-known measure in financial analysis is the cost-revenue ratio, often viewed as a measure of financial efficiency. Here we modify it by inversion to obtain the revenue-cost ratio (E), which therefore is also a measure of efficiency.

$$(1)\quad E = \frac{TR}{TC}$$

To achieve an operating profit, revenue must exceed cost so the value of E must be greater than one. The greater the ratio, the greater the profit. Otherwise, when E equals one, the company is breaking even, while if E is less than one an operating loss is present.

Financial professionals often find it useful to consider percentages, not just ratios, when discussing the relationship between cost and revenue. To this end, we can derive specific implications for two measures of the profitability rate. One commonly used measure, *the profitability percentage* is the amount of profit divided by the total cost of producing and selling that output.

The profitability rate can be determined by subtracting from E the value of one and converting to a percentage (e.g., if E =1.1, the profitability rate is ten percent). As the value of efficiency (E) rises, so, too, does the profitability rate.

To show this more formally, suppose profit is π dollars, the premium of total revenue over total cost (π = TR - TC). Equation (1) can be expanded as:

$$(1a) \quad E = \frac{TR}{TC} = \frac{TC}{TC} + \frac{\pi}{TC} = 1 + \pi/TC$$

Since a value of one alone for E (the first term on the far right above) implies neither profit nor loss (break-even), the last term on the right (π/TC) is a measure of profitability which can be converted into percentages by multiplying by 100. That is, the profitability percentage (pp) is defined as:

$$(2) \quad pp = \left(\frac{\pi}{TC}\right) x\ 100 = (\ E - 1)\ x\ 100$$

A second measure of profitability, the *profitability margin* (pm below), is the ratio of profit (π) to total revenue. The exact relationship between the profitability margin and efficiency (E) is:

$$(3) \quad pm = \frac{(TR-TC)}{TR} = 1 - \frac{TC}{TR} = 1 - 1/E$$

Equation (3) implies the greater the value of E at any level of output, the higher the profitability margin. Qualitatively, it is not surprising that a greater value of E results in a higher profitability margin, but the precise amount is identified mathematically by the above equation.

For example, if E =1 in the above, pm will equal zero, which follows from recognizing that in this case, total revenue must equal total cost. But if E rises to 1.05, equation (3) implies that the profitability margin will be about 4.8%.

Indeed, the extent to which *raising* E will *increase* the profitability margin in general can be determined mathematically from the equation above to be:

$$(3a) \quad \frac{d\ pm}{d\ E} = (1/E)^2$$

In words, the profitability margin will increase in a nonlinear fashion (i.e., at a diminishing rate) as E increases. Other measures of rates of profitability can also be related and defined to multifactor productivity in a manner analogous to that above. Table 1 below shows relationships among E, pp and pm.

Table 1 - Quantifying relationships among E, pp and pm.

		Table 1		
	Relationship Between E, pp and pm			
	RC	pp	pm	Δpm*
	1.00	0.0%	0.0%	
	1.05	5.0%	4.8%	4.8%
	1.10	10.0%	9.1%	4.3%
	1.15	15.0%	13.0%	4.0%
	1.20	20.0%	16.7%	3.6%
	1.25	25.0%	20.0%	3.3%
	1.30	30.0%	23.1%	3.1%
	1.35	35.0%	25.9%	2.8%
	1.40	40.0%	28.6%	2.6%
	1.45	45.0%	31.0%	2.5%
	1.50	50.0%	33.3%	2.3%
	1.55	55.0%	35.5%	2.2%
	1.60	60.0%	37.5%	2.0%
	1.65	65.0%	39.4%	1.9%
	1.70	70.0%	41.2%	1.8%
	1.75	75.0%	42.9%	1.7%
	1.80	80.0%	44.4%	1.6%
	1.85	85.0%	45.9%	1.5%
	1.90	90.0%	47.4%	1.4%
	1.95	95.0%	48.7%	1.3%
	2.00	100.0%	50.0%	1.3%
	Δpm is the percentage point change			
	in pm that results from an increment			
	of .05 in E			

The efficiency ratio E is also consistent with microeconomics. There, efficiency is often described as minimum average (or unit) cost. Expanding the definition of E by representing total revenue as the product of average revenue (p) and quantity, we obtain:

(4) $E = TR/TC = p*Q/AC*Q = p*(1/AC)$ or p/AC

This implies the greatest efficiency (value of E) is achieved when AC is minimized, which corresponds to the minimum point on the average cost curve.

Microeconomic theory, however, implies in the short run (when fixed costs are present) the greatest efficiency (operating at the lowest unit cost) is usually not the most profitable level of output. This is because total profit equals the profit margin multiplied by the level of output. A lower profit margin may be more than offset by a greater volume of output.

A simpler explanation of the link between average cost and profit follows using Cost-Volume-Profit (CVP) Analysis, which focuses on determining the output level needed to avoid a loss (break-even) or achieve a target profit. In this framework, the formula (derived in the Appendix) determining the break-even level of output (Q_{BE}) is:

(5) $Q_{BE} = FC/(p-v)$

where FC represents fixed cost, p is the price (average revenue) of a unit of output, and v is the variable cost per unit of output (assumed constant).

To achieve a target level of profit (T), the output level (Q_T) must necessarily be higher. The formula becomes:

(6) $Q_T = (FC + T)/(p\text{-}v)$

A reduction in average cost (efficiency) can be achieved through either a lower fixed cost or variable cost per unit. If either occurs, the formulas above imply the break-even and target profit levels of output are reduced. Alternately expressed, a higher target can be achieved at any given level of output. So profit targets set to comply with a conservatively lower required quantity of output can now be raised.

The linkage between greater efficiency (lower average cost) and profitability using the CVP model is depicted in the chart below. The curves illustrate profit (vertical axis) at each volume (horizontal axis) for two cost situations. In the first, profit (π) is calculated based on a fixed cost of $100 and a variable cost per unit of $10, while for the second (π') fixed cost is $50 and variable cost per unit is $7. The selling price of the product is $20. Since the curve associated with the first cost set has a lower vertical intercept and smaller slope than the second it is easier (lower volume needed) to break-even or achieve a specified target profit. Decision-makers can decide on a target profit based upon risk preference and anticipated demand.

Linkage to productivity

At paper's outset it was noted 'Some companies or industries have experienced impressive productivity gains over time without a commensurate gain in earnings (e.g. airlines).' This finding refers to 'labor productivity,' the most widely used measure. But is this the case when more than one input is introduced into the measure, in what is termed 'multifactor productivity?'

Multifactor productivity has been defined in several ways, often depending upon the number of inputs, from the simplest case considering two inputs, to more advanced ones considering several, sometimes many more inputs (see the measure developed at the U.S. Bureau of Labor Statistics). In each case, a challenge is how to combine the inputs in the denominator of the multifactor productivity measure. There are issues in how to measure each input, how to combine inputs often expressed in different units, and what equation specification is most appropriate in identifying multifactor productivity. While addressed in advanced studies, the complexity, time and effort needed to construct the multifactor measure for an organization limits the utility of the concept as a practical tool in business decision-making.

Some Production and Operations Management texts suggest a simplified version that preserves the measure's comprehensive nature while expediting application by reducing data requirements. This version is useful in identifying how multifactor productivity in an organization has changed over time or differs among competitors at a given time. It is perhaps not useful in directly identifying the most efficient mix of inputs. But by making comparisons over time and across industry, can offer some suggestions as to what the most useful combination and use of inputs may be.

In this managerial accounting-type framework, input usage is not expressed in different units, but at cost. Each input could be weighted and then combined, or simply added if the motivating factor is to focus

on total cost efficiency rather than the productive relationship itself. Two versions of a multifactor productivity (MFP) measure then follow, differing in terms of how the numerator is defined. The first is:

$$(7a)\quad MFP_1 = Q/TC$$

Where Q, as before, is the number of units of output produced and TC designates the total factor cost associated with input use. The inverse of the expression on the right is average cost (AC), thereby linking this version of MFP to cost efficiency described earlier.

When units of output are heterogeneous, output is often represented by sales revenue. If so, the multifactor productivity measure becomes:

$$(7b)\quad MFP_2 = TR/TC$$

which is the revenue/cost ratio addressed earlier in the paper.

So MFP_2 is also a measure of efficiency, but one that has a direct link to profitability rates and levels as shown earlier. The higher the value of MFP_2 the greater cost effectiveness and profit ability.

Applying the concepts to actual data – the Commercial airline industry

One industry for which the underlying data components are available is the U.S. Airline industry, which reports information on traffic, revenue and cost to the Federal Aviation Administration (FAA). Industry evidence based upon these data which has been organized and presented by the MIT Airline Data Project (ADP) is now used to evaluate relationships among multifactor productivity, efficiency, and profitability.

Besides the two profitability rate measures used earlier, the analysis now explores the relationship of actual MFP for the industry as a whole and for two sectors to the widely used measure of unit profit.

Converting expressions for output and cost above into airline industry terms.

A common measure of number of output units produced is "available seat miles" (ASM), the aggregate number of seats on multiplied by the distance traveled on each flight.
Total value of output unit is based on fares plus revenue received from fees such as seat selection, baggage, and priority boarding, and transport of goods charges. Average total revenue per available seat mile, called TRASM is equal to total **operating revenue** divided by ASM, serves as the measure of unit market value. Thus, MFP is:

(8a) $MFP_2 = (ASM \times TRASM)/TC$

In a similar fashion, total cost is the mathematical product of cost per available seat mile (CASM) and the number of ASM, so:

(8b) $MFP_2 = (ASM \times TRASM)/(ASM \times CASM)$, or simplifying

(8) $MFP_2 = TRASM/CASM$

CASM is regarded by industry analysts as a measure of airline efficiency, specifically "cost-efficiency."

Notice by replacing the right-hand side of equation (7a) by the inverse of CASM, MFP_1 can be viewed as a measure of productivity and cost-efficiency.

A useful distinction, then. between the two multifactor productivity measures is that MFP_1 focuses on cost-efficiency, while MFP_2 considers the effect of cost and revenue efficiency (TRASM) in a combined manner.

Results obtained from using the MIT Airline Data Project (ADP) data for 2006-2015 to calculate each MFP and profitability measure examined above are shown in Table 2.

There are three panels in the table, the first for all domestic airlines (collectively) and then separate panels for two sub- categories of airlines, network carriers and low-cost carriers, which have often been viewed as differing in productivity and profitability. The traditional view is that LCCs have consistently been more profitable than network carriers because of lower costs.

MFP and Profit for Domestic Carriers – Using MIT ADP Data

Total All Sectors	2007	2009	2011	2013	2015
MFP_2 (op rev/op exp)	1.0529	1.0012	1.0357	1.0763	1.2074
pp	5.29%	0.12%	3.57%	7.63%	20.74%
pm	5.03%	0.12%	3.45%	7.09%	17.18%
MFP_1 = ASM/oper. expense	7.86	7.67	6.53	6.45	7.40
operating income per ASM (unit profit)	0.0067	0.0002	0.0055	0.0118	0.0280
network carriers	2007	2009	2011	2013	2015
MFP_2(op rev/op exp)	1.051	0.987	1.031	1.069	1.191
pp	5.13%	-1.26%	3.15%	6.95%	19.09%
pm	4.88%	-1.28%	3.05%	6.50%	16.03%
operating income per ASM (unit profit)	0.0070	-0.0018	0.0052	0.0117	0.0282
LCC carriers					
MFP_2 (op rev/op exp)	1.068	1.043	1.036	1.076	1.244
pp	6.85%	4.28%	3.57%	7.60%	24.44%
pm	6.41%	4.10%	3.45%	7.07%	19.64%
operating income per ASM (unit profit)	0.0062	0.0042	0.0044	0.0093	0.0264

Table 2

The table demonstrates the similarity in movement between MFP_2 and all three profitability measures. In years in which productivity has increased so, commensurately, has profitability. When productivity has fallen, profitability has, too.
Indeed, the correlation between MFP_2 and pm and pp is perfect (i.e. an identity), while that between MFP_2 and unit profit is in excess of 99%. Improving MFP_2, therefore, is synonymous with higher profitability evidenced in all three measures.

Concluding Remarks

Multifactor productivity is an efficiency measure, but as shown above, it has a clear and direct relationship to profitability. Specifically, the analysis above demonstrated that two profitability measures, profitability margin and profitability percentage, bear a direct, exact linear mathematical relationship with MFP_1. Also shown was the mathematical connection between MFP_1 and unit profit, which is nonlinear. To now examine empirically how closely the behavior of unit profit by year varies with MFP_1, I have calculated the correlation coefficient (linear) between these variables. The result, which appears in line 3 of Table 3 below, shows that the relationship is very close to perfectly linear over the range of data found in this study. So the linkage between MFP_1 and all three profitability measures is direct and very strong.

What about the relationship between MFP_2 and each of these profitability measures? The analysis above indicated that the connection depends upon the level of TRASM, and so is less compelling. To see the empirical connection, I have calculated correlation coefficients between MFP_2 and each profitability measure. Each correlation is very low. It appears that MFP_2 is far more telling as a cost-efficiency measure than as a link to profitability because it does not consider revenue-efficiency.

Much more widely used as an efficiency measure, of course, is labor productivity. It is a partial measure since it considers only one input, but it could be useful to determine the extent to which movements in labor productivity coincide with movements in multifactor productivity and profitability. A usual measure of labor productivity, units of output divided by units of labor (FTEs), have been computed from ADP data.

These correlations are reasonably high, suggesting that improved labor productivity can lead to greater profitability, but nowhere as compelling as the mathematical and empirical linkages between MFP_1 and the profitability measures.[i]

REFERENCES

Bingham, Eugene and Houston, Joseph. Fundamentals of Financial Management, 15th edition, Cengage, 2018

Thomas, Christopher and Maurice, S. Charles. Managerial Economics. 13th edition, McGraw Hill, 2020.

MIT Airline Data Project – Data on Domestic Airlines

Stevenson, Wayne, Operations Management, 14th edition, McGraw-Hill/Irwin, 2021.

Appendix Table 1 – Productivity, Output and Corporate Profits from Domestic Sources, 2006-2012

Source: U.S. Bureau of Labor Statistics

Year	labor prod	business output	corporate profits
2006 QTR 1	2.5	3.5	**1,344.0**
2006 QTR 2	0.5	3.5	**1,342.4**
2006 QTR 3	-2.5	2.4	**1,404.1**
2006 QTR 4	2.3	2.8	**1,315.4**
2007 QTR 1	-0.5	1.2	**1,227.5**
2007 QTR 2	4.2	1.9	**1,252.8**
2007 QTR 3	4.8	2.8	**1,127.1**
2007 QTR 4	1.5	2.1	**1,031.6**
2008 QTR 1	-2.2	1.4	**941.9**
2008 QTR 2	2.5	0.4	**915.0**
2008 QTR 3	-0.9	-1.8	**889.2**
2008 QTR 4	-3.1	-5.2	**621.1**
2009 QTR 1	5.3	-5.6	**850.2**

2009 QTR 2	7.1	-6.2	**921.8**
2009 QTR 3	5.8	-4.8	**1,040.1**
2009 QTR 4	5.0	-0.5	**1,146.1**
2010 QTR 1	2.2	1.6	**1,259.3**
2010 QTR 2	-0.6	2.9	**1,217.1**
2010 QTR 3	3.2	3.9	**1,343.1**
2010 QTR 4	1.5	3.5	**1,384.0**
2011 QTR 1	-2.5	2.6	**1,302.7**
2011 QTR 2	1.1	2.6	**1,364.7**
2011 QTR 3	0.5	1.9	**1,382.7**
2011 QTR 4	2.9	2.4	**1,502.2**
2012 QTR 1	-0.6	3.2	**1,497.2**
2012 QTR 2	1.7	2.9	**1,485.3**

[i] The coefficient of determination (r^2), i.e. the explained variation in profitability as related to variation in productivity is much lower for labor productivity (in the 50-60% range) compared to the nearly perfect relationship of profitability with MFP_1.

Big Data, Analytics, and Knowledge Management

A STRATIFICATION-BASED CLUSTERING ANALYSIS APPROACH TO PERFORMANCE EVALUATION OF DECISION-MAKING UNITS

TRACK: Big Data, Analytics, and Knowledge Management

Competitive Paper Submitted for Presentation at the 52nd NEDSI Annual Conference
March 30-April 1, 2023

A STRATIFICATION-BASED CLUSTERING ANALYSIS APPROACH TO PERFORMANCE EVALUATION OF DECISION-MAKING UNITS

ABSTRACT

Since data envelopment analysis (DEA) was introduced, DEA-based methods have emerged as a popular technique for evaluating the performance of decision-making units (DMUs) and ranking them in diverse areas. In addition, DEA has evolved into a tool for Big Data analysis. The critical issue of these models is that each DMU's unfavorable inputs/outputs are ignored to maximize its own efficiency. As a result, these methods lack discriminating power and consequently generate inconsistent rankings, especially for efficient DMUs. In addition, the DEA-based methods require impractical computation time for a large set of DMUs. This paper proposes a new stratification-based clustering analysis (SBCA) method to eliminate such critical issues. The proposed method demonstrates its outstanding performance through numerical examples, let alone removing the crucial shortcomings of DEA-based methods, which include the excessive computation time for solving a large set of DMUs in the Big Data context.

KEYWORDS: Data envelopment analysis, Decision-making unit, Discriminating power, Stratification, Clustering analysis

INTRODUCTION

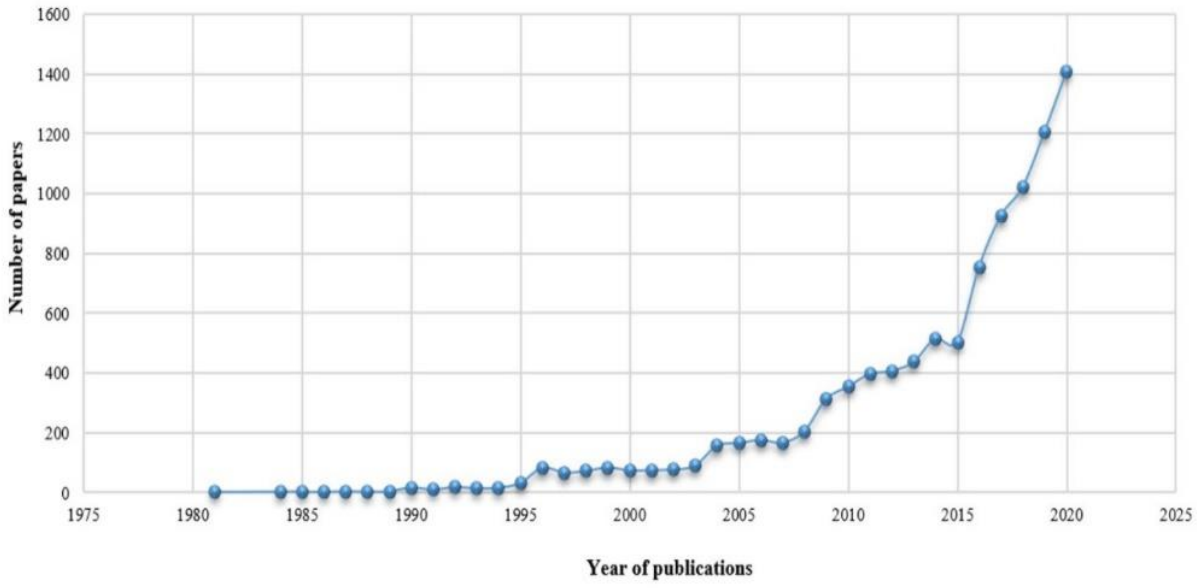
Among many performance assessment methods, Data envelopment analysis (DEA) has emerged as one of the most popular techniques and has evolved into a tool for Big Data analysis. DEA uses Linear Programming (LP) to rate the relative performance of a set of peer decision-making units (DMUs) by comparing how well the DMU converts s its inputs into its outputs. As the whole technique is based on comparing each DMU with the remaining ones, a considerably large number of DMUs compared to the number of outputs and inputs are necessary for the evaluation to be meaningful (Ramanathan, 2003; Meza & Jeong, 2013). The traditional DEA (T-DEA), introduced by Charnes et al. (1978), eventually determines which DMUs attain the most efficient outcome using given inputs and which do not. The T-DEA model generates a single, comprehensive performance measure, called an efficiency score (ES), for each DMUs.

Consequently, the T-DEA model classifies all DMUs into two groups, which would separate efficient units from inefficient units. For the inefficient units, the analysis can quantify

what levels of improved performance should be achievable and what levels of enhanced performance should be achievable. The analysis can also indicate where the inefficient DMU might search for benchmarking as it searches for ways to improve. Also, it suggests where such an inefficient DMU might look for benchmarking help as it pursues ways to improve.

Performance measurement or evaluation is one of the most emphasized elements in any performance management. Hence, T-DEA's empirical orientation and absence of prior assumptions make DEA-based models one of the most popular tools for assessing and improving operational processes. As Zhu (2022) points out, the various methods based on the T-DEA model have been developed and widely applied in many different areas to evaluate the relative performance of DMUs, resulting in **more than 5,000 publications in the Web of Science database**. Panwar et al. (2022) also observe that there has been a constant growth in the publications of DEA from 1978 to 1995. Still, from 1995 onwards, there has been an exponential rise in theoretical development and diverse applications. They show such a trend in Figure 1.

Figure 1. DEA publications since 1980



Three comprehensive review papers (Bhat et al., 2019; Cui and Yu, 2021; Panwar et al., 2022) show how widely DEA has been applied even in sports and the airline industry. Zhu

(2022) suggests that DEA should be regarded as a method or tool for data-oriented analytics in performance evaluation and benchmarking. However, the DEA methods have shown some drawbacks. T-DEA frequently finds too many efficient DMUs, out of all DMUs under evaluation. Thus, there is no way to rank efficient DMUs since this method can't distinguish efficient DMUs.

To remedy this deficiency of T-DEA, Sexton et al. (1986) suggest the cross-evaluation DEA (CE-DEA) method to evaluate and rank DMUs, with the main idea of using T-DEA to do the peer evaluation rather than T-DEA's pure self-evaluation. The CE-DEA method can usually provide a complete ranking of the assessment of DMUs (see Anderson et al., 1993). As a result of its enhanced discriminating power, the CE evaluation has found a significant number of applications in the DEA literature (see Liang et al., 2008; Wang and Chin, 2010; Zerafat Angiz et al., 2013; Gavgani and Zohrehbandian, 2014; Hou et al., 2018; Lee, 2019; Liu et al., 2019). As Doyle and Green (1994) pointed out, the non-uniqueness of CE scores (CESs) often results from alternative optimal weights in the T-DEA model, implying the CESs depend on the optimization software used. Thus, the non-uniqueness has been criticized as a major drawback of using the CE method as a ranking tool.

The ranking for DMUs would imply substantial meanings since decision-makers can be aware of the current level of a DMU's performance. They can establish incentive schemes, support policies, and develop strategies to sustain a future business or core competence. Thus, the ranking of DMUs can be used to identify competitors with similar performance levels and determine the future direction of efficiency improvement for sustainable development (Lee and Choi, 2019).

Each DEA-based method has its own critical issues, drawbacks, or disadvantages, which could become severe as the number of inputs or outputs increases. As said before, any DMUs with at least one specialized input and output can obtain an efficient status. In addition, the ES of a specific DMU depends upon the other DMUs in the reference set. Thus, DEA-based models frequently produce inconsistent ESs and unstable rankings for the under-assessment DMUs. For example, when all DMUs are evaluated, the top-ranked DMU should maintain the top-ranking position regardless of the cases when some lower-ranked DMUs are not assessed together. But the DEA-based models frequently allow the previously lower-ranked DMUs to overtake the higher-ranked DMUs if some lower-ranking DMUs are not evaluated together. In addition, the

CE-DEA method often ranks inefficient DMUs higher than the efficient DMUs identified by T-DEA. These weaknesses would be unavoidable because these DEA-based models have unfairly treated unfavorable inputs/outputs. These critical issues make it difficult for decision-makers to evaluate and rank DMUs impartially. No literature has explicitly and seriously discussed these essential weaknesses of the DEA-based models regarding unstable and unreasonable rankings generated by these models.

Cluster analysis, a data mining technique, is an exploratory data analysis tool to find meaningful groups whose members have something in common. It aims to sort entities into groups or clusters so that the degree of association is robust among members of the same cluster and feeble among members of different clusters (see Aggarwal, 2015; Delen, 2021). One of many tools for calculating distance measures is the centroid-based technique. This paper proposes a stratification-based clustering analysis (**SBCA**) technique utilizing the overall minimum, average, and maximum points to represent three clusters rather than the centroid of each cluster. If a DMU's inputs are closer to the minimum point than the average point, it would belong to the efficient group, say cluster #1, regarding its inputs. Similarly, a DMU will belong to the efficient group, cluster #1, regarding its output if its outputs are closer to the overall maximum point than the average point. The DMUs, which do not belong to either efficient or inefficient, would be classified into cluster #2. Based on the clusters to which each DMU's inputs and outputs belong, all DMU are stratified into one of three levels. Without ignoring or sacrificing any given inputs/outputs, the SBCA method would eliminate the critical weaknesses of DEA-based evaluation models and rate the DMUs more impartially and consistently than the DEA-based methods.

The DEA-based methods have recently been viewed as a tool for Big Data-enabled analytics. The critical issue is that DEA-based methods require excessive computation time to solve a large set of DMUs in the Big Data. The proposed process takes advantage of the simple computation and gets the results very quickly, not requiring any optimization software needed by all DEA-based models. Therefore, the proposed SBCA method could help practitioners and researchers perform a more refined and consistent efficiency evaluation than the DEA-based methods. Besides, the proposed procedure would quickly provide a benchmarking framework for DMUs to improve efficiency due to the cluster level they belong.

LITERATURE REVIEW

Several review papers on the DEA-based ranking methods are available as an attempt to close the gap in the literature on the ranking methods in the DEA context, such as Adler et al. (2002), Jahanshahloo et al. (2008), Markovits-Somogyi (2011), Lotfi et al. (2013), and Aldamak and Zolfaghari (2017). Table 1 lists these review papers chronologically (See Hong, 2021). The DEA methods or models that each paper reviews and each paper's major findings are also briefly explained. These review papers reveal that there have existed a variety of articles that apply different ranking methods, seeking to improve the discriminative power of DEA methods and to rank all DMUs fully.

Table 1 reveals that the typical ranking methods/models are cross-efficiency, super-efficiency, benchmarking, multi-criteria decision-making, and statistical analysis. Adler et al. (2002) insist that many mathematical and statistical techniques have been presented to increase the DEA-based methods' discriminating power and rank the DMUs effectively. But, while each method may be suitable in a specific area, no one methodology can be agreed upon as all ills' panacea. They (2002) expect the decisive DEA model to be developed to eliminate all weaknesses or problems and be easy and convenient to solve by practitioners in the field and academics alike. After 15 years since they (2002) expected the decisive DEA model to be developed, Aldamak and Zolfaghari (2017) review one hundred and twelve (112) articles published in various scholarly journals on the subject of DEA-based ranking methods. They (2017) observe that none of the proposed DEA-based ranking categories is optimum for every evaluation assessment, asserting that no ranking method has been proven to be either an ideal or widespread way for ranking the efficiency of DEA-based models. They (2017) also insist that the absence of global evaluation criteria makes it impossible to evaluate all the presented methods reviewed in their paper. Depending on the evaluation's nature, they (2017) conclude that each approach could be better than others according to the decision maker's preferences and evaluation objectives. Panwar et al. (2022), reviewing two hundred and sixty-six (226) articles on the main subject of historical development and current trends of DEA models, publish the recent and most comprehensive review paper on the DEA. They do not include DEA’s ranking method but agree with what previous review papers have found by reiterating the drawbacks or disadvantages of each DEA model.

Table 1. The review papers on the DEA-based ranking methods

Authors	Title	# of Reviewed References	Reviewed Ranking Models	Summary
Adler et al. (2002)	"Review of ranking methods in the data envelopment analysis context"	59	Benchmarking, Super efficiency, Cross efficiency, Proportional measures, Multivariate statistics, Multi-criteria decision-making method	Several mathematical and statistical techniques have been presented to increase DEA's discriminating power and fully rank the DMUs under evaluation
Jahanshahloo et al. (2008)	"Review of ranking models in the data envelopment analysis"	22	Using a common set of weights, Super efficiency-based AP, MAJ, Revised MAJ, Slack adjusted, Gradient line models, L1 Norm, Concept of advantage model, Tchebycheff Norm., Monte Carlo method, Slack-based model (SBM).	Some models are infeasible for unique data. The Monte Carlo method is suggested to rank all efficient DMUs because it ranks extreme and non-extreme DMUs.
Markovits-Somogyi (2011)	"Ranking efficient and inefficient decision-making units in data envelopment analysis"	43	Benchmarking, Super efficiency, Cross efficiency, Minimum weight restriction, Statistical analysis, Slack-based DEA, Application of fuzzy logic, Shadow prices, Multi-criteria decision-making methods,	The full ranking is achievable through several ways in DEA, and the choice of a specific method would depend on the particular needs of the study in question.
Lotfi et al. (2013)	"A review of ranking models in data envelopment analysis"	104	Benchmarking, Super efficiency, Cross efficiency, Optimal weights in DEA, Multivariate Statistics, Stratification, Gradient line, Multi-criteria decision-making method.,	The DEA ranking is reviewed and classified into seven general groups. Some reviewed models are applied to the example, and none of the DMUs is ranked consistently.
Aldamak and Zolfaghari (2017)	"Review of efficiency ranking methods in data envelopment analysis"	112	Benchmarking, Super efficiency, Cross efficiency, Statistics, Common weights, Inefficient frontier, Virtual DMU, DM interference, Fuzzy concept, Multi-criteria decision-making method	The absence of universal assessment criteria makes it impossible to evaluate the presented methods. According to the decision-maker's preferences and evaluation objectives, each method could be better than the others.

As mentioned before, each method discussed in the previous section has shown its own weaknesses. Several DEA-based models sometimes rank efficient DMUs differently (Hong and Jeong, 2017). Decision-makers are usually interested in selecting the top-ranked DMU(s) before choosing the efficient ones among DMUs. If the #1 ranked DMUs depend on the evaluation method, it will confuse the decision-makers or practitioners. The proposed method could be better than the DEA-based methods for answering the age-old question, 'which DEA ranking method should we use?'.

Some researchers apply a DEA-based approach for clustering. In other words, they build clusters for each DMU based on the results of DEA rather than a distance measure. For example, Po et al. (2009) propose a DEA-based clustering approach rather than a distance measure. Their proposed method employs the piecewise production function derived from the T-DEA method to group the input and output items. Following Po et al. (2009), Chen et al. (2022) propose a clustering approach within the CE-DEA framework to obtain a clustering result based on the production relationship between the inputs and outputs date of DMUs. But these DEA-based clustering methods would not be free from DEA's intrinsic weakness of ignoring unfavorable inputs and outputs.

DATA ENVELOPMENT ANALYSIS-BASED MODELS

The ratio-form or fractional DEA model uses the ratio of outputs to inputs to measure the relative efficiency of DMU_j as an objective function to be assessed relative to the ratios of all DMUs. The fractional DEA model is stated (Zhu, 2014):

Objective Function: Maximize the efficiency rating θ for DMU_j

$$Max\ \theta_j\ = \frac{\sum_{r=1}^{s} u_{rj} y_{rj}}{\sum_{i=1}^{m} v_{ij} x_{ij}}, \quad j = 1, 2, \dots n, \tag{1}$$

which is subject to the constraint that when the same set of u and v multipliers is applied to all other DMUs to be compared, all DMUs will be less than or equal to 100% efficient as follows:

$$\frac{\sum_{r=1}^{s} u_{rj} y_{rw}}{\sum_{i=1}^{m} v_{ij} x_{iw}} \leq 1, \forall j\ \&\ w = 1, 2, \dots, n. \tag{2}$$

where

j = DMU j being evaluated in the DEA analysis, $j = 1,\ldots, n$

y_{rj} = amount of output r generated by DMU_j

x_{ij} = amount of input i used by DMU_j

m = number of inputs used by DMUs

s = number of outputs generated by DMUs

u_{rj} = multipliers or weight assigned by DEA to output r for DMU_j

v_{ij} = multipliers or weight assigned by DEA to input i for DMU_j

$u_{rj}, v_{ij} \geq 0, r = 1, \ldots, s; i = 1, \ldots, m; w = 1, 2, \ldots, n.$

Based on the fractional DEA model in (1)-(2), the traditional (T) and cross-efficiency (CE) DEA models were developed (see Cooper et al., 2011). The T-DEA model is formulated as the following LP problem, where E_{jj} represents the efficiency score (ES) for DMU_j:

$$max \quad E_{jj} = \sum_{r=1}^{s} u_{rj} y_{rj}, \tag{3}$$

subject to

$$\sum_{i=1}^{m} v_{ij} x_{ij} = 1, \tag{4}$$

$$\sum_{r=1}^{s} u_{rw} y_{rw} - \sum_{i=1}^{m} v_{ij} x_{iw} \leq 0, w = 1, \ldots, n, \tag{5}$$

$$u_{rj}, v_{ij} \geq 0, r = 1, \ldots, s; i = 1, \ldots, m.$$

The CE-DEA method, which consists of two phases, was proposed to rank DMUs with the central idea of using DEA to do peer assessment rather than pure self-assessment (see Sexton et al., 1986). The weights or multipliers from the first phase are applied to all DMUs to get the cross-efficiency score (CES) for each DMU in the second phase. In the first phase, the above LP model in (3)-(5) is solved to find the ES of DMU_j. To denote the peer evaluation, let E_{jw} represent the DEA score for the rated DMU_w, w = 1, 2, …, n, using the optimal weights /multipliers that a rating DMU_j has chosen in the model (3)-(5). Now, E_{jw} is given by

$$E_{jw} = \frac{\sum_{r=1}^{s} u_{rj}^{*} y_{rw}}{\sum_{i=1}^{m} v_{ij}^{*} x_{iw}}, \qquad j \; and \; w = 1, \ldots, n. \tag{6}$$

Then, the CE score for DMU_w is defined as follows:

$$CE_w = \frac{1}{n}\sum_{j=1}^{n} E_{jw}. \quad (7)$$

STRATIFICATION-BASED CLUSTERING ANALYSIS (SBCA) METHOD

The proposed approach, which evaluates each DMU's performance without applying DEA methods, differs from some DEA-based clustering methods. In the proposed SBCA method, there are three (3) clusters to that a DMU would belong. The first cluster is for efficient DMUs, the second is for less efficient DMUs, and the third is for inefficient DMUs. All DMUs under evaluation will be stratified into one of these three clusters for input and output, respectively. Rather than the centroid-based clustering method, the SBCA method utilizes the overall minimum, average, and maximum points to represent these three clusters. Efficiency or productivity is expressed as the ratio of outputs to inputs. It implies that as an output gets greater and/or an input gets less, the greater the efficiency will be.

Let c_{iI}^{k} and c_{rO}^{k} denote the representatives of the k^{th} cluster for input (*I*) *i* and output (*O*) *r*, respectively, where *k* =1, 2, and 3. Then we define

$$c_{iI}^{1} = Min_{\forall j}\{x_{ij}\}, \quad (8)$$

$$c_{rO}^{1} = Max_{\forall j}\{y_{rj}\}, \quad (9)$$

$$c_{iI}^{2} = \{\bar{x}_{ij}\}, \forall j, \quad (10)$$

$$c_{rO}^{2} = \{\bar{y}_{rj}\}, \forall j, \quad (11)$$

$$c_{iI}^{3} = Max_{\forall j}\{x_{ij}\}, \quad (12)$$

$$c_{rO}^{3} = Min_{\forall j}\{y_{rj}\}, \quad (13)$$

Note that c_{iI}^{1} and c_{iI}^{3} are the overall minimum and maximum points for input *i*, c_{rO}^{1} and c_{rO}^{3} are the overall maximum and minimum points for output *r*, and c_{iI}^{2} and c_{rO}^{2} are the average points for input *i* and output *r*, respectively. Now, let α_i and β_r denote the weights assigned to input *i* and output *r*, respectively. The more (less) important an input or an output is, the greater (less) α_i or

β_r is applied. Then, the weighted Euclidean distances (WEDs) for DMU_j from the representatives of these two clusters are computed as follows:

(i) For input (*I*),

$$WED_{jI}^{k} = \sqrt{\sum_{i=1}^{m} \alpha_i (x_{ij} - c_{iI}^{k})^2}, \forall k. \quad (14)$$

Let μ be the cluster # that satisfies the following equation:

$$WED_{jI}^{\mu} = Min_{\forall k}\{WED_{jI}^{k}\}, \ \forall \mu. \quad (15)$$

Then, we classify DMU_j into the cluster, C_I^{μ}, μ = 1, 2, and 3, in terms of input (*I*).

(ii) For outputs (*O*),

$$WED_{jO}^{k} = \sqrt{\sum_{r=1}^{s} \beta_r (y_{rj} - c_{rO}^{k})^2}, \forall k. \quad (16)$$

Let λ be the cluster # that satisfies the following equation:

$$WED_{jO}^{\lambda} = Min_{\forall k}\{WED_{jO}^{k}\}, \ \forall \lambda \quad (17)$$

Similarly to the case of input, we classify DMU_j into the cluster, C_O^{λ}, λ = 1, 2, and 3, in terms of output (*O*). Then, there will be nine cases for the combination of clusters where DMU_j belongs, in terms of input and output, respectively. We assign five (5) stratification points (S-Pts) if a DMU belongs to the first cluster for each C_I^1 and C_O^1, three (3) S-Pts for each second cluster, C_I^2 and C_O^2 and one (1) S-Pt for each third cluster, C_I^3 and C_O^3. Table 2 presents the nine cases, the corresponding S-Pt for each case, and the total S-Pts. Based on the total S-Pts, DMUs are stratified into three efficiency levels, ℓ = 1, 2, and 3. Level 1 consists of DMUs with a total S-Pts of 8 or above, DMUs with a total S-Pts of 6 are the members of Level 2, and Level 3 comprises DMUs with a total S-Pts of 4 or less. Note that each Level encompasses three cases. Level 1 includes cases {#1, #2, #4}; Level 2 {#3, #5, #7}; and Level 3 cases {#6, #8, and #9}.

Table 2. Combinations of clusters for input and output and stratification points

No.	Input (I)	S-Pt for Input (I)	Output (O)	S-Pt for Output (O)	Total S-Point	Stratification (Level, ℓ)
1	C_I^1	5	C_O^1	5	10	1
2	C_I^1	5	C_O^2	3	8	1
3	C_I^1	5	C_O^3	1	6	2
4	C_I^2	3	C_O^1	5	8	1
5	C_I^2	3	C_O^2	3	6	2
6	C_I^2	3	C_O^3	1	4	3
7	C_I^3	1	C_O^1	5	6	2
8	C_I^3	1	C_O^2	3	4	3
9	C_I^3	1	C_O^3	1	2	3

Once all DMUs are stratified into one of three levels, the next step is to find the efficiency scores (ESs) for ranking DMUs under evaluation. The efficiency is expressed as the ratio of desired output to required input. We temporarily define ES as the Euclidean distance (ED) ratio of outputs from the overall minimum point to the ED of inputs from the overall minimum point. Thus, the greater output and/or the lower input, the greater the efficiency. Now, the unnormalized efficiency score of DMU_j, ES_j^U, is given by

$$ES_j^U = \frac{ED\ of\ outputs\ of\ DMU_j\ from\ the\ minimum\ of\ outputs}{ED\ of\ inputs\ of\ DMU_j\ from\ the\ minimum\ of\ inputs}. \quad (18)$$

ES_j^U in (18) increases as the numerator increases and/or the denominator decreases. Note that WED_{jI}^1 in (14) becomes the denominator of the formula of ES_j^U. But WED_{jO}^3 in (16) should not be the numerator in (18) because a DMU with a longer WED_{jO}^3, which denotes the WED from the overall minimum point, is not necessarily more efficient than a DMU with a shorter WED_{jO}^1, which denotes the WED from the overall maximum point. For the numerator of ES_j^U in (18), we define the total WED (TWED) for output as the WED between the overall maximum and minimum points. Then, $TWED_O$, which denotes the TWED for output (O), is given by

$$TWED_O = \sqrt{\sum_{r=1}^{s} \beta_r (c_{rO}^1 - c_{rO}^3)^2}, \quad (19)$$

where c_{rO}^{1} in (9) and c_{rO}^{3} in (13) are the representatives of the 1st cluster and 3rd cluster for output (*O*) *r* of DMUs, respectively. Then, WED_{jO}^{3R}, which denotes the revised WED_{jO}^{3} for DMU_j, is obtained by subtracting WED_{jO}^{1} in (16) from $TWED_O$ in (19) as follows:

$$WED_{jO}^{3R} = TWED_O - WED_{jO}^{1}. \quad (20)$$

Now, ES_j^U in (18) should be expressed as the ratio of WED_{jO}^{3R} in (20) to WED_{jI}^{1} in (14), as follows:

$$ES_j^U = \frac{WED_{jO}^{3R}}{WED_{jI}^{1}}. \quad (21)$$

If $WED_{jI}^{1}$is zero or too small, then set $WED_{jI}^{1} < Min_{\forall w}\{WED_{wI}^{1}\}, \forall w \neq j$. If $WED_{jO}^{3R} = 0$, recompute $TWED_O$ in (19) by setting $c_{rO}^{3} = \sigma_r * Min_j\{y_{rj}\}$, where $0 \leq \sigma_r < 1$ to prevent ES_j^U from becoming zero. The normalized ES, ES_j^N, for DMU_j in level ℓ would be obtained by diving the maximum unnormalized one as follows:

$$ES_j^N = \frac{WED_{jO}^{3R}/WED_{jI}^{1}}{Max_{\forall \omega}\{WED_{\omega O}^{3R}/WED_{\omega I}^{1}\}}. \quad (22)$$

We rank DMUs based on the normalized ES, ES_j^N. As long as the ranks generated by the normalized ES do not change, ES_j^N in (22) can be adjusted by increasing WED_{jI}^{1} of the most efficient DMU, i.e., $ES_j^N = 1$.

NUMERICAL EXAMPLES

First, we consider the numerical example Liang et al. (2008) illustrate with five DMUs. The data for each DMU consists of three inputs denoted by (x_{1j}, x_{2j}, x_{3j}) and two outputs represented by (y_{1j}, y_{2j}), as shown in Table 3. Table 4 reports ES and CES generated by T- and CE-DEA and also reports ES_j^N by SBCA, along with the corresponding rank for each DMU. From Table 4, we note that DMU_3, one of the two efficient DMUs by T-DEA, is a dominating unit regarding all three inputs and one output (y_{2k}) and is ranked as #1 by all three methods, T-, CE-DEA, and SBCA method. The other efficient DMU_2 is ranked #2 by CE-DEA and SBCA methods. As the DEA models implicitly assume equal weight to each input and output, it is unreasonable for

the DEA-based models to rank *DMU₄* higher than *DMU₅*. We observe that *DMU₄* and *DMU₅* have the same minimum input, 5.0, and the maximum output, 6.0. But *DMU₄* has the other two inputs, 9.0 and 8.0, which are greater than or equal to those of *DMU₅*, 6.0, and 8.0. Except for the same maximum output value of 6.0, *DMU₅* has an output value of 3.0, higher than 2.0 for *DMU₄*, implying that *DMU₅* should be ranked higher than *DMU₄*. Thus, **the ranks generated by the SBCA method would be more rational and logical than those by CE-DEA.**

After stratification, we report results for the two efficient DMUs, *DMU₃* and *DMU₂*, in Table 5. For comparison purposes, the original ranks before stratification, denoted by R_{BF}, based on the ranks in Table 4. Table 4 exhibits that SBCA consistently ranks two efficient *DMU₃* and *DMU₂* as #1 and #2, but the top-ranked DMU is changed from *DMU₃* to *DMU₂* by CE-DEA after stratification. Table 5 also reports various absolute rank differences (ARDs) to investigate each method's robustness in generating rankings. From ARD^2 and ARD^3 between ranks with and without stratification, we can see that the rankings generated by SBCA are consistent, whereas CE-DEA generates inconsistent rankings. For comparison between CE-DEA and SBCA, we report ARD^4, which shows the discrepancy in ranks between these two methods.

Table 3: Five decision-making units

DMU	Input			Output	
	x_{1j}	x_{2j}	x_{3j}	y_{1j}	y_{2j}
1	7.0	7.0	7.0	4.0	4.0
2	5.0	9.0	7.0	7.0	7.0
3	4.0	6.0	5.0	5.0	7.0
4	5.0	9.0	8.0	6.0	2.0
5	6.0	8.0	5.0	3.0	6.0

Table 4: Comparison of efficiency scores and rankings before stratification

DMU	T-DEA		CE-DEA		SBCA				
	ES	R^1	CES	R^2	Input (I)	Output (O)	Level (ℓ)	ES_j^N	R^3
1	0.6857	5	0.5191	4	C_I^2	C_O^2	2	0.3278	4
2*	1.0000*	1	0.9161	2	C_I^2	C_O^1	1	0.9716	2
3*	1.0000*	1	0.9571	1	C_I^1	C_O^2	1	1.0000	1
4	0.8571	3	0.6985	3	C_I^3	C_O^3	3	0.1698	5
5	0.8571	3	0.4942	5	C_I^2	C_O^2	2	0.4576	3

*: Efficient by T-DEA; R: Rank;

Table 5: Comparison of efficiency scores and rankings after stratification

DMU	CE-DEA				Level (ℓ)	SBCA				ARD^4
	CES	R^2	R^2_{BF}	ARD^2		ES^N_j	R^3	R^3_{BF}	ARD^3	
2*	1.000	1	2	1	1	0.836	2	2	0	1
3*	0.946	2	1	1	1	1.000	1	1	0	1

*: Efficient by T-DEA; R_{BF}: Rank before stratification; ARD^2= |R^2-R^2_{BF}|; ARD^3= |R^3-R^3_{BF}|; ARD^4 =|R^2- R^3|.

To further investigate the performance of the SBCA method, we consider the data of Zhu (2014), which is presented in Table 6. There were fifteen (15) companies on the Top Fortune Global 500 list in 1995, with three inputs: (i) assets ($ millions), (ii) equity ($ millions), and (iii) number of employees, and with two outputs: (i) revenue ($ millions) and (ii) profit ($ millions). In Table 7, we report ES generated by T-DEA, CES by CE-DEA, and ES^N_j by the proposed SBCA method, along with the corresponding ranks. There are seven (7) efficient companies with a perfect ES of 1.000 by T-DEA and six (6) DMUs classified into level 1 (ℓ =1) by SBCA. The CE-DEA model ranks *DMU$_5$* 'Sumitomo' and *DMU$_9$* 'Exxon' as #1 and #2 efficient companies, respectively. In contrast, the SBCA method agrees with CE-DEA regarding #1 company *DMU$_5$* 'Sumitomo', but ranks *DMU$_3$* 'Itochu' as #2. For the efficient DMUs, the ranks by CE-DEA and SBCA are compatible except for *DMU$_{11}$* 'Walmart' and *DMU$_9$* 'Exxon.' As pointed out as a critical weakness, the CE-DEA method ranks two efficient *DMU$_4$* 'General Motors' and *DMU$_{13}$* 'Nippon Life' lower than some inefficient DMUs, such as *DMU$_6$* 'Marubeni' and *DMU$_{10}$* 'Royal Dutch/Shell', ranked #5 and #7.

In fact, *DMU$_{13}$* 'Nippon Life', rated as efficient by T-DEA, is ranked #11 and #14 by CE-DEA and SBCA, respectively. It exhibits a critical weakness of T-DEA, which allows each DMU to ignore unfavorable inputs and/or outputs. From Table 6, 'Nippon Life' has the largest 'Assets,' but the smallest 'Equity' and relatively small 'Employees' as its inputs. 'Nippon Life' does not show any potential of being rated as efficient in terms of outputs. But, T-DEA allows this unit to ignore its unfavorable input 'Assets' and output 'Revenue.' Obviously, the other two inputs, 'Equity' and 'Employees,' and an output, 'Profit,' are considered for 'Nippon Life' to be rated as efficient by T-DEA. *DMU$_4$* 'General Motors', one of the two DMUs ranked lower than some inefficient DMUs, shows a similar weakness of T-DEA. Being rated #8 and #10 by CE-DEA and SBCA, *DMU$_4$* 'General Motors' with the highest 'Employees' and the other two high inputs is rated as efficient due to the two high outputs.

Table 6: Fifteen (15) companies from the Fortune Global 500 List of 1995

DMU	Company	Input			Output	
		Assets	Equity	Employees	Revenue	Profit
1	Mitsubishi	91,920.6	10,950.0	36,000	184,365.2	346.2
2	Mitsui	68,770.9	5,553.9	80,000	181,518.7	314.8
3	Itochu	65,708.9	4,271.1	7,182	169,164.6	121.2
4	General Motors	217,123.4	23,345.5	709,000	168,828.6	6,880.7
5	Sumitomo	50,268.9	6,681.0	6,193	167,530.7	210.5
6	Marubeni	71,439.3	5,239.1	6,702	161,057.4	156.6
7	Ford Motor	243,283.0	24,547.0	346,990	137,137.0	4,139
8	Toyota Motor	106,004.2	49,691.6	146,855	111,052.0	2,662.4
9	Exxon	91,296.0	40,436.0	82,000	110,009.0	6,470.0
10	Royal Dutch/Shell	118,011.6	58,986.4	104,000	109,833.7	6,904.6
11	Wal-Mart	37,871.0	14,762.0	675,000	93,627.0	2,740.0
12	Hitachi	91,620.9	29,907.2	331,852	84,167.1	1,468.8
13	Nippon Life	364,762.5	2,241.9	89,690	83,206.7	2,426.6
14	Nippon T & T	127,077.3	42,240.1	231,400	81,937.2	2,209.1
15	AT&T	88,884.0	17,274.0	299,300	79,609.0	139.0

We collectively select all seven efficient DMUs by T-DEA and level 1 DMUs by SBCA and report the results for the nine DMUs chosen in Table 8. As described in the first example, the rankings, R_{BF}^2 and R_{BF}^3, denote the expected ranks based on the ranks in Table 7, where all DMUs are evaluated. For example, 'General Motors' is ranked #8 and #10 by CE-DEA and SBCA, respectively, as shown in Table 7. Out of nine (9) DMUs in Table 8, the expected rank for 'General Motors' is #7 by both methods. Note that five DMUs, {Mitsui, Itochu, Sumitomo, Marubeni, Exxon}, are not only efficient but also are classified into level 1. Table 8 shows that the SBCA method consistently ranks 'Sumitomo' and 'Itochu' as the #1 and #2 DMUs. But, CE-DEA switches the ranks for the top-two DMUs, 'Sumitomo' and 'Exxon, so 'Exxon' is now #1, and 'Sumitomo' is #2, showing the critical weakness of CE-DEA. DMU_6 'Marubeni', identified as inefficient by T-DEA, is still ranked higher than efficient DMUs, {Mitsui, Wal-Mart, Nippon Life}, by CE-DEA. ARD^3 in Table 8 clearly shows that the SBCA method generates robust rankings for all efficient and level 1 DMUs, whereas CE-DEA generates inconsistent rankings, as ARD^2 indicates.

Table 7. Comparison of diverse efficiency scores and rankings before stratification

DMU	Company	T-DEA		CE-DEA		Input (I)	Output (O)	Level (ℓ)	SBCA		ARD
		ES	R^1	CES	R^2				ES_j^N	R^3	
1	Mitsubishi	0.6628	11	0.5727	9	C_I^1	C_O^1	1	0.4908	4	5
2	Mitsui	1.0000*	1	0.7850	4	C_I^1	C_O^1	1	0.3796	5	1
3	Itochu	1.0000*	1	0.8795	3	C_I^1	C_O^1	1	0.9831	2	1
4	General Motors	1.0000*	1	0.6806	8	C_I^3	C_O^1	2	0.0383	10	2
5	Sumitomo	1.0000*	1	0.9334	1	C_I^1	C_O^1	1	1.0000	1	0
6	Marubeni	0.9720	8	0.7723	5	C_I^1	C_O^1	1	0.7443	3	2
7	Ford Motor	0.7372	10	0.4653	10	C_I^2	C_O^2	2	0.0450	9	1
8	Toyota Motor	0.5246	12	0.3788	12	C_I^2	C_O^2	2	0.0600	8	4
9	Exxon	1.0000*	1	0.8970	2	C_I^1	C_O^2	1	0.0949	6	4
10	Royal Dutch/Shell	0.8414	9	0.6911	7	C_I^2	C_O^2	2	0.0683	7	0
11	Wal-Mart	1.0000*	1	0.7513	6	C_I^3	C_O^3	3	0.0066	11	5
12	Hitachi	0.3861	13	0.2954	13	C_I^2	C_O^3	3	0.0043	12	1
13	Nippon Life	1.0000*	1	0.4354	11	C_I^2	C_O^3	3	0.0034	14	13
14	Nippon T & Ta	0.3486	14	0.2863	14	C_I^2	C_O^3	3	0.0031	13	1
15	AT&T	0.2704	15	0.1828	15	C_I^2	C_O^3	3	0.0001	15	0

ARD(Absolute Rank Difference) = $|R^2 - R^3|$

Table 8. Comparison of rankings for the top nine DMUs selected collectively after stratification

DMU	Company	T-DEA		CE-DEA				Level (ℓ)	SBCA				ARD^4
		ES	R^1	CES	R^2	R^2_{BF}	ARD^2		ES^N_j	R^3	R^3_{BF}	ARD^3	
1	Mitsubishi	0.6628	9	0.5138	9	8	1	1	0.4751	4	4	0	5
2	Mitsui	1.0000	1	0.6696	6	4	2	1	0.3674	5	5	0	1
3	Itochu	1.0000	1	0.8732	3	3	0	1	0.9474	2	2	0	1
4	General Motors	1.0000	1	0.7578	5	7	2	2	0.0369	7	7	0	2
5	Sumitomo	1.0000	1	0.8763	2	1	1	1	1.0000	1	1	0	1
6	Marubeni	0.9719	8	0.7782	4	5	1	1	0.7144	3	3	0	1
9	Exxon	1.0000	1	0.9781	1	2	1	1	0.0841	6	6	0	5
11	Wal-Mart	1.0000	1	0.6645	7	6	1	3	0.0049	8	8	0	1
13	Nippon Life	1.0000	1	0.6055	8	9	1	3	0.0001	9	9	0	1

Level: Level from Table 7; ARD^2 (Absolute Rank Difference) = $|R^2-R^2_{BF}|$; ARD^3 = $|R^3-R^3_{BF}|$; ARD^4 = $|R^2- R^3|$

For further comparison, we select the top five DMUs collectively. These selected DMUs are evaluated, and evaluation results are shown in Table 9. Out of the five chosen DMUs, two DMUs {Mitsubishi, Marubeni} are inefficient and three {Itochu, Sumitomo, Exxon} are efficient by T-DEA, while SBCA classifies all five DMUs into level 1. CE-DEA still ranks 'Exxon' as #1, while SBCA ranks it as the bottom, #5. The proposed SBCA approach generates consistent rankings again with the sum of ARD^2 of zero.

In Table 10, we list the top four DMUs by each method for each case, where all 15 DMUs are evaluated in Case 1, all efficient DMUs and level 1 DMUs are assessed in Case 2, and the top-four DMUs in Case 2 are evaluated. Specifically, 'Sumitomo' and 'Itochu" are consistently ranked among the top four DMUs by both CE-DEA and SBCA. 'Marubeni,' inefficient by T-DEA but classified into level 1 by SBCA, ranked five times as one of the top four. A notable observation is that 'Exxon,' ranked #1 or #2 by CE-DEA, is ranked #6 twice by SBCA.

SUMMARY AND CONCLUSIONS

Various DEA-based methods have enabled decision-makers to measure the efficiency of DMUs and rank them based on efficiency scores (ESs). Ranking DMUs based on the ESs generated by the T-DEA method is meaningless since it can only separate efficient DMUs from inefficient ones. In addition, T-DEA shows a significant drawback because of its self-evaluation principle. Several ranking methods based on the T-DEA have been proposed, but no method has been found to be either a universal or superior method for ranking DMUs. The authors, who developed various DEA-based ranking methods, have asserted that the absence of global assessment criteria makes it impossible to evaluate all the presented methods reviewed by their papers. They conclude that each method could be better than others according to the decision maker's preferences and evaluation objectives, depending on the evaluation's nature (see Aldamak and Zolfaghari, 2017).

The DEA-based approaches' severe weakness is their biased preference for specific inputs/outputs. All DMUs under evaluation can only use favorable inputs/outputs to boost their own ESs, dropping the unfavorable inputs/outputs. This paper proposes the stratification-based clustering analysis (**SBCA**) method to evaluate DMUs more consistently without prejudice, using the overall minimum, average, and maximum values to represent three clusters rather than the centroid of each cluster. We apply the proposed SBCA method to evaluate the two well-

known numerical examples, which have been considered by several authors, to compare the DEA-based methods. These numerical examples show that the rankings generated by T-DEA or CE-DEA show such a significant weakness, especially for the top-notched DMUs. As shown in the numerical example, when the top-notched DMUs are evaluated, a previously lower-ranked DMU takes over the top-ranked unit. Some inefficient DMUs are ranked higher than several efficient DMUs. By contrast, the proposed SBCA generates consistent rankings for the level 1 DMUs through numerical examples, even when removing some lower-level, lower-ranked, or inefficient DMUs from evaluation. In the numerical examples, we assume equal weights, $\alpha_i = 1, \forall i$, for inputs and $\beta_r = 1, \forall r$, for outputs. Another advantage of the SBCA method is that we can evaluate and rank the DMUs more flexibly by allowing the decision-maker to assign unequal weights to each input or output.

The DEA-oriented methods in the Big Data context have been developed and proposed. Barr and Durchholz (1997) show that as the number of DMUs increases, the DEA methods require impractical amounts of computation/run time. Several models are proposed to reduce the computation time necessary for solving large-scale DEA problems without evaluating all DMUs (Khezrimotlagh et al., 2019). Contrary to the DEA-oriented methods, the SBCAmethod does not require any optimization software, and it takes a short computation time to get the evaluation results. A popular DEA software, DEAFrontier, is run on Intel® Xeon ® Gold 5122 HP Z4 Workstation PC (2 processors) with 32GB of RAM installed using a 64-bit version of Windows 10. **The computation time for finding the CESs for the second numerical example with fifteen (15) DMUs turns out to be 10.33 sec. It takes only 0.12 sec for the SBCA method with VBA (Visual Basic for Applications) code in Excel to evaluate the same DMUs.** We randomly generate 200 DMUs with three inputs and two outputs for the experimental purpose and compare the computation time. It takes **258.61 sec** for CE-DEA but only **1.18 sec** for SBCA, showing a significant advantage for the proposed method regarding computation time.

The results and observations through the numerical examples demonstrate that the proposed method would be considered an appropriate tool for evaluating a large set of DMUs. In fact, the SBCA method could be used as an essential tool, along with various DEA-based methods, for decision-makers to rate and rank DMUs under evaluation in the Big Data context. Future research would be interesting and necessary to apply the SBCA method extensively in the Big-Data context and real-world DEA applications.

Table 9. Comparison of rankings for the top five DMUs selected collectively after stratification

DMU	Company	T-DEA		CE-DEA				Level (ℓ)	SBCA				ARD^4
		ES	R^1	CES	R^2	R^2_{BF}	ARD^2		ES^N_j	R^3	R^3_{BF}	ARD^3	
1	Mitsubishi	0.6628	5	0.4282	5	5	0	1	0.3277	4	4	0	1
3	Itochu	1.0000	1	0.9428	2	3	1	1	0.9286	2	2	0	0
5	Sumitomo	1.0000	1	0.9428	2	1	1	1	1.0000	1	1	0	1
6	Marubeni	0.9719	4	0.8858	4	4	0	1	0.5883	3	3	0	1
9	Exxon	1.0000	1	1.0000	1	2	1	1	0.0007	5	5	0	4

Level: Level from Table 7; ARD^2 (Absolute Rank Difference) = |R^2-R^2_{BF}|; ARD^3 = |R^3-R^3_{BF}|; ARD^4 = |R^2- R^3|

Table 10. Top four companies for each case

	CE-DEA			SBCA		
Rank	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
1	Sumitomo	Exxon	Exxon	Sumitomo	Sumitomo	Sumitomo
2	Exxon	Sumitomo	Sumitomo	Itochu	Itochu	Itochu
3	Itochu	Itochu	Itochu	Marubeni	Marubeni	Marubeni
4	Mitsui	Marubeni	Marubeni	Mitsubishi	Mitsubishi	Mitsubishi

Case 1: All DMUs are evaluated before stratification.
Case 2: Top 8 DMUs after stratification are evaluated.
Case 3: Top 4 DMUs after stratification are evaluated.

REFERENCES

Adler, N., Friedman, L., & Sinuany-Stern, Z. (2002). Review of ranking methods in the data envelopment analysis context. *European Journal of Operational Research*, *140*, 249-265.

Aggarwal, C. (2015). *Data mining: The textbook*. Springer, New York.

Aldamak, A., & Zolfaghari, S. (2017). Review of efficiency ranking method in data envelopment analysis. *Measurement*, *106*, 161-172.

Andersen, P. M., & Petersen, N.C. (1993). A Procedure for ranking efficient units in Data Envelopment Analysis. *Management Science*, *39*, 1261-1264.

Barr R. S., & Durchholz, M. (1997). Parallel and hierarchical decomposition approaches for solving large-scale Data Envelopment Analysis models. *Annals of Operations Research*, *73*, 339-372.

Bhat, Z.U.H., Sultana, D., & Dar, Q.F. (2019). A comprehensive review of data envelopment analysis (DEA) in sports. *Journal of Sports Economics & Management*, *9*(2), 82-109.

Charnes, A., Cooper, W.W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operations Research*, *2*(6), 429-444.

Chen, L., Wang, S.H., & Wang, Y.M. (2022). A new data envelopment analysis clustering approach within cross-efficiency framework. *Journal of the Operational Research Society*, *73*(2), 664-673.

Cooper, W.W., Seiford, L.M., & Zhu, J. (2011). *Data envelopment analysis: history, models and interpretations*, In: Cooper et al. (Ed.), Handbook on data envelopment analysis, (pp. 1-39), New York: Springer.

Cui, Q., & Yu, L-T. (2021). A review of data envelopment analysis in airline efficiency: State of the art and prospects. *Journal of Advanced Transportation*, Article ID2931734.

Delen, D. (2021). *Predictive analytics: Data mining, machine learning, and data science for practitioners*. 2nd Edition, Pearson Education, New Jersey.

Doyle, J., & Green, R. (1994). Efficiency and cross efficiency in DEA: derivations, meanings and uses. *Journal of the Operational Research Society*, *45*(5), 567-578.

Gavgani, S.S., & Zohrehbandian, M. (2014). A cross-efficiency-based ranking method for finding the most efficient DMU. *Mathematical Problems in Engineering*, 2014, Article ID. 269768.

Hong, J.D. (2021). Multiple criteria DEA-based ranking approach with the transformation of decision-making units. *International Journal of Applied Industrial Engineering*, *8*(1), 1-20.

Hong, J.D., & Jeong, K.Y. (2017). Cross-efficiency-based heuristics to rank decision-making units in data envelopment analysis. *Computers & Industrial Engineering*, *111*, 320-330.

Hou, Q., Wang, M., & Zhou, X. (2018). Improved DEA cross-efficiency evaluation method based on ideal and anti-ideal point. *Discrete Dynamics in Nature and Science*, 2018, Article ID. 1604298.

Jahanshahloo, G.R., Lotfi F.H., Sanei, M., & Jelodar M.F. (2008). A review of ranking models in data envelopment analysis. *Applied Mathematical Sciences, 2*(29), 1431-1448.

Khezrimotlagh, D., Zhu, J., Cook, W.D., & Toloo, M. (2019). Data envelopment analysis and big data. *European Journal of Operational Research*, 274, 1047-1054.

Lee, Y.C. (2019). Ranking DMUs by combining cross-efficiency scores based on Shannon's entropy. *Entropy*, *21*(5), 467.

Lee, J., & Choi, G. (2019). A dominance-based network method for ranking efficient decision-making units in data envelopment analysis. *Sustainability*, *11*(7): 2059.

Liang, L., Wu, J., Cook, W., & Zhu, J. (2008). The DEA game cross-efficiency model and its nash equilibrium. *Operations Research*, *56*, 1278-1288.

Liu, H.H., Song, Y.Y., & Yang, G.L. (2019). Cross-efficiency evaluation in data envelopment analysis based on prospect theory. *European Journal of Operational Research*, *273*(1), 364-375.

Lotfi, F.H., Jahanshahloo, G.R., Khodabakhshi, J., Rostamy-Malkhlifeh, M., Moghaddas, Z., & Vaez-Ghasemi, M. (2013). A review of ranking models in data envelopment analysis. *Journal of Applied Mathematics*, 2013, Article ID 492421.

Markovits-Somogyi, R. (2011). Ranking efficient and inefficient decision-making units in data envelopment analysis. *International Journal for Traffic and Transportation Engineering*, *1*(4), 245-256.

Meza, D., & Jeong, K. (2013). Measuring efficiency of lean six sigma project implementation using data envelopment analysis at NASA. *Journal of Industrial Engineering Management*, *6*(2), 401-422.

Panwar, A., Olfati, M., Pant M., & Snalsel, V. (2022). A review on the 40 years of existence of data envelopment analysis models: Historic development and current trends. *Archives of Computational Methods in Engineering*, 29, 5397-5426.

Po, R. W., Guh, Y.Y., & Yang, M.S. (2009). A new clustering approach using data envelopment analysis. *European Journal of Operational Research*, *199*, 276-84.

Ramanathan, R. (2003). *An introduction to data envelopment analysis – a tool for performance measurement*. New Delhi, India: Sage Publications.

Sexton, T. R., Silkman, R.H., & Hogan, A.J. (1986). Data envelopment analysis: critique and extensions, In: Silkman, R.H. (Ed.), *Measuring efficiency: An assessment of data envelopment analysis* (32, pp. 73-105), San Francisco: Jossey-Bass.

Wang, Y.M., & Chin, K.S. (2010). A neutral DEA model for cross-efficiency evaluation and its extension. *Expert Systems with Applications*, *37*, 3666-3675.

Zerafat Angiz, M., Mustafa, A., & Kamali, M.J. (2013). Cross-ranking of decision-making units in data envelopment analysis. *Applied Mathematical Modelling*, *37*, 398-405.

Zhu, J. (2014). *Quantitative models for performance evaluation and benchmarking: data envelopment analysis with spreadsheets*. 3rd Edition, Springer, New York.

Zhu, J. (2022). DEA under big data: data-enabled analytics and network data envelopment analysis. *Annals of Operations Research*, *309*(2), 761-783.

IS MONEYBALL EFFICIENT?

A DEA Approach to Analyzing Baseball

Abstract

Analytics has transformed the game of baseball. But is it an efficient approach to playing the game? In this paper we propose a data envelopment analysis model to determine the relative efficiency with which players play the game. Since the goal of winning requires scoring runs, this model demonstrates an approach in calculating the relative efficiency of players turning plate appearances into getting on base. Expanding upon previous work, the dual model is also implemented for analyzing the inputs to determine where and by how much a player may be inefficient.

KEYWORDS: Baseball, Data Envelopment Analysis, moneyball, sabermetrics, linear programming, analytics

Introduction

Baseball has always been a game of statistics. Data has been collected on every at bat going back a century. From this, batting averages, and earned run averages became the standards for comparison. Current changes in how these statistics are being used to play the game is the new key to analytics or what is euphemistically called *moneyball*.

Before the onset of analytics, teams scouted high school and college players, and if the scouts thought that they met a series of criteria they would sign them to a contract. At they same time they would spent millions signing free agents who had already demonstrated their abilities in the major leagues. The complaint of the A's front office was that as a small market team they could not afford the salaries of the stars and would thus be locked out of long term competitiveness(Lewis, 2003). But late in the twentieth century, the Oakland A's shook up professional baseball by changing how teams were being managed.

At that time the A's tried a different approach - moneyball. The theory was that since pennants are won by maximizing the number of wins, a team should be focusing on what creates wins. In its most basic form wins require runs which require runners to reach base. If instead of looking for sluggers you instead focus on players who hit singles or have a high number of walks, the team could increase wins without the exorbitant cost of home run hitters. Retired player and then A's general manager, Billy Beane, took the idea to new heights in setting the A's roster, falling just short of the pennant in 2002. The success of moneyball in Oakland expanded the use of analytics in baseball. Data collection and analysis is now implemented in every aspect of the game.

History of statistics in baseball

It would be wrong to assume that analyzing baseball began with the Oakland A's. Statistics has been a fundamental component of the game since Henry Chadwick's development of the box score in 1858 (Chadwick, 1872). A game in which measurements are overwhelmingly discrete; number of hits, number of at bats, number of strike outs, ... is paradise for a statistician. From these you get batting averages, slugging averages, on base percentages, and earned run averages. Computer analysis has not changed the process, it has simply eased their use. In the late 1960s, Davey Johnson, All Star second baseman for the Baltimore Orioles and a mathematics student in college would provide Hall

of Fame manager Earl Weaver with his analyses titled *Optimize Orioles Lineup*. Earl would promptly toss the printout in the waste basket (Wulf, 2012). But if you knew Earl Weaver and his approach to the game, you would know that he had his own form of analytics. He just did his analysis on note cards instead of spreadsheets.

The modern approach to baseball statistics began with Earnshaw Cook and *Percentage Baseball* (Cook, 1964), but took off with the rise of *sabermetrics* and Bill James' *Baseball Abstracts* (James, 1982). First shunned by the leagues, it slowly took foot and has become what is commonly called analytics.

In the analytics approach to baseball, data is analyzed not only at the level of the individual game, but also to the batter and pitch. This includes pitches thrown, bases on balls, singles, stolen bases, and many more. The problem may be that statistics are intended to provide a big picture; a means of cutting through the noise in order to make decisions. Instead, analytics may have become micromanaging. As an example, a pitcher might be pulled based on their pitch count instead of how they are throwing that day. Or a batter is told to take or swing away based upon the pitcher, runners on base, and the current count of balls and strikes. Taking this down to the minute details, analytics has created a game in which each individual pitch and swing can be determined by the numbers.

In an odd way, analytics may have upended the entire concept of moneyball; using statistics to find underrated and thus underpaid players. The degree of analytical decision making in the individual games has become the goal. Today's stars may be stars because of analytics, but finding underrated talent may still not be there. As Wharton professor Abraham Wyner states *What we assume that the teams should know, but never seem to get, is that you're paying for the future not the past* (Wyner, 2019).

Applying data envelopment analysis to baseball

An alternative to studying the minutia is to use a more broad analytical criteria to find efficient players that may be overlooked. Since the goal is to get runners on base, in the model the runners need to be the output. The inputs can be the players themselves, but instead of isolated measurements such as hits, and runs scored, as a single means of production. In this way the relative efficiency of the player can be determined. This approach can be performed using Data Envelopment Analysis.

In DEA, the goal is to determine the relative efficiency of an individual decision making unit with respect to other DMU's in the same market. Doing so is performed by determining how well a set of similar inputs is used in producing one or more outputs. Using linear programming, a type of linear regression, not only can the efficiencies be calculated, but the locations of the inefficiencies determined as well.

Data Envelopment Analysis has been applied to baseball in the past. Mazur (Mazur, 1994) usedan individual player was determined to be the decision making unit. The inputs for the model were number of home runs, and number of runs batted in for an individual player over a single season while the single output being the player's batting average.

A second application of DEA takes the same approach that each player is a single DMU. But in this analysis the production input is at bats while the outputs are getting on base by counting the number of singles, double, triples, home runs, and base on balls (Anderson & Sharp, 1997).

These applications all have the same goal of determining the relative efficiency of players, but none aplply this to the sabermetric ideas touted in moneyball. That is that the goal is that to win games batters must get on base. This paper applies data envelopment analysis to this application; determining the relative efficiency of a batter turning plate appearances into getting on base.

Methodology

In production economics, a common question is how efficient is the production unit? The type of production group can be as varied as the industry in which they function. In production analysis each group might be assumed to operate independently of the others. This independence can be described in terms of decision making authority for their own operations. As such it is common to call each of these groups a *decision making unit* or *DMU*. In banking the DMUs might be the individual branches. In a hospital they could be the individual departments. In a university the separate colleges.

This efficiency analysis is on major league baseball. Thus the DMU could be a league, a team, or an individual player. At the player level, we would need to differentiate between pitchers and batters, analyzing them separately. For this model we will limit ourselves to batters, although a similar analysis can be applied to pitchers as well.

Development of the DEA Model

In its most general form efficiency is defined as the ratio of output y to input x. If we let efficiency be represented by η, then

$$\eta_{\text{abs}} = \frac{y}{x} \tag{1}$$

In batting, we can define x as the number of at bats while y is the number of hits. In this case the efficiency is just the common *batting average* statistic . Batting average is actually an absolute efficiency in that it is independent of any other player and thus can be used to compare batters both within and outside of any specific group.

As long as the absolute efficiencies are compared within a single industry the values are comparable. But it would be misleading if there was an attempt to compare the efficiency of a player in one league to one in another. After all, an all star major league player with a $.340$ batting average is not the same as the high school player batting $.650$.

Since efficiencies should only be compared within the industry, it can be useful to make them relative to to others within their group. In baseball, the best batter in a season would be assigned the relative efficiency of 1.000 with all of the others correspondingly scaled to it. In keeping with the multivariate efficiencies to come, to scale the absolute efficiency to relative efficiency it is necessary to multiply the input and output by a fixed weight.

$$\eta_k = \frac{\nu}{\mu}\frac{y_k}{x_k} \tag{2}$$

where

$$\frac{\nu}{\mu} = \max_{k=1\ldots n}\left(\frac{x_k}{y_k}\right)$$

In the case of batting averages ν/μ is simply the reciprocal of the maximum batting average in the analysis. This would result in earning the league leader a relative efficiency of 1.0 with the rest of the players scaled from that accordingly.

If every batter had the same skill set, that is they hit the same percentage of singles, doubles, triples, and home runs, then a single measure of efficiency would be fine. But they do not. There are career single hitters, and home run sluggers. And while they may both have the same number of plate appearances they will be making different contributions to the game. As a result, it is necessary to define a relative efficiency that may have multiple inputs and multiple outputs.

The same approach as in equation 2 can be extended to two inputs and two inputs. The relative efficiency for player

k is

$$\eta_k = \frac{\nu_1 y_{1k} + \nu_2 y_{2k}}{\mu_1 x_{1k} + \mu_2 x_{2k}} \tag{3}$$

There is no reason that the number of inputs and outputs must be equal, so assume that m is the number of inputs and n the number of outputs. Further, assume that there are N DMUs. The relative efficiency for player k is

$$\eta_k = \frac{\sum_{i=1}^{n} \nu_i y_{jk}}{\sum_{i=1}^{m} \mu_h x_{ik}} \tag{4}$$

The multivariate relative efficiency in equation 4 requires a means of estimating the weights for each input and output variable. Since relative efficiencies are constrained such that $0 \leq \eta_k \leq 1$, the relative efficiency for each DMU becomes a constraint. For DMU k,

$$0 \leq \frac{\sum_{i=1}^{n} \nu_i y_{jk}}{\sum_{i=1}^{m} \mu_h x_{ik}} \leq 1 \tag{5}$$

or

$$\sum_{i=1}^{n} \nu_j y_{ik} \leq \sum_{i=1}^{m} \mu_h x_{ik}$$

$$\sum_{i=1}^{n} \nu_j y_{ik} - \sum_{i=1}^{m} \mu_h x_{ik} \leq 0 \tag{6}$$

$$\tag{7}$$

A further constraint on the relative efficiency is that the output for the current DMU's outputs be unity.

$$\sum_{i=1}^{m} \mu_i x_{ik} = 1 \tag{8}$$

To maximize the outputs for that same DMU, this becomes the linear program that is known as the CCR data envelopment model(Ramanathan, 2003). Again, for DMU k,

$$\begin{aligned}
&\max z = \sum_{i=1}^{n} \nu_i y_{ik} - \epsilon \sum_{i=1}^{m} x_{ik} \\
&\text{subject to} \\
&\sum_{i=1}^{m} \mu_i x_{ik} = 1 \\
&\sum_{i=1}^{n} \nu_i y_{ij} - \sum_{i=1}^{m} \mu_h x_{ij} \leq 0 \qquad\qquad j = 1, 2, \ldots, N \\
&x_{ij}, y_{ij}, \epsilon > 0
\end{aligned} \tag{9}$$

In this ϵ is some small non-zero value. This linear program would then be repeated for each $k = 1 \ldots N$.

The linear program in equation 9 will estimate the relative efficiency of each DMU. But it will also provide estimates of the slack for each output variable. If the DMU is efficient, $\eta_k = 1$, the slacks will all be zero. But if $\eta_k < 1$ the DMU is inefficient. In this case the slack will be difference between the current output and the amount that the output variable would need to be increased to achieve efficiency,

In addition to the primal model to estimate the relative efficiency of each DMU there is the dual model. The dual in

this type of analysis is of great importance in that the slack for each variable provides the amount that the input would need to be reduced to achieve efficiency relative to the efficient DMUs.

$$\begin{aligned}
&\min z = \sum_{i=1}^{N} \Theta_j \\
&\text{subject to} \\
&\sum_{i=1}^{N} x_{ij}\Theta_j \geq x_{ik} && j = 1, 2, \ldots m \\
&\sum_{i=1}^{N} y_{ij}\Theta_j > 0 && j = 1, 2, \ldots n
\end{aligned} \tag{10}$$

As before, k indicates the individual DMU being analyzed.

Application of DEA to baseball analysis

The outputs of a DEA model are the results of the production process while the inputs are the raw materials. For the baseball analysis, previous research defined output as the number of hits, singles, doubles, home runs, and base on balls, and the number of at bats as a single input(Anderson & Sharp, 1997). This is a direct extension of the traditional batting average. Because of an oddity of the way that baseball statistics are calculated, official at bats do not include bases on balls or sacrifice outs. As such an alternative is to look at plate appearances. This variable is a count of the number of times that a batter comes up to bat regardless of the outcome. Plate appearances are given the variable PA

The choice of outputs in hitting analysis is straightforward. The goal is to put runners on base while not making an out. Thus the obvious outputs are the number of times that the batter successfully reaches base. This is the number of singles, doubles, home runs, and base on balls that a batter has over some time period. In this analysis they will be indicated as 1B, 2B, HR, and BB. There is a question that should be asked about not including triples in this mix. The exclusion was intentional. The number of triples for most players over an entire season is in the low single digits if not zero. Since a top player will make over 500 plate appearances in a season triples, while exciting, are considered insignificant in the majority of games.

There are additional variables that while they might be an output in a production sense, are actually detrimental to the goal of getting on base, scoring runs, and winning games. In a single word - outs. Each time a batter approaches the plate there is an exceptionally high probability that they will fail, and instead of getting on base they will make an out. Each game has a finite number of outs, 27 in a standard nine inning game, and once reached the game is over. While making an out is an output of the production, it is an undesirable output.

Undesirable outputs have been addressed in the literature(Jabhanshahloo, Lotfi, Shoja, Razavyan, & Tohidi, 2005). A simple approach to addressing this is to treat an undesirable output as an input. The theory is that since the goal of any production system is maximizing the output of the production process while minimizing the inputs to the process, if an output is undesirable then it should be minimized as well. An example from manufacturing is pollution. Pollution is the output of the manufacturing process, but since the goal is to minimize this output, it might be included as an input in a DEA model.

In baseball, making an out is clearly an undesirable output. Even in the case of a sacrifice bunt or a sacrifice fly which moves a runner into scoring position or scores a runner from third respectively, doing the same without making an out would still be preferred. Because of this these types of outs are not included in the analysis. But other outs are even more undesirable. Two in particular can be significant; strikeouts and fly outs. In keeping with the standard

scoring nomenclature, K and F. Clearly striking out is undesirable output. An out is recorded with no beneficial result. A similar result is the fly out. It often results from an attempt at hitting a home run. Again, an out is recorded and no beneficial result occurs. A fly out is not to be confused with a sacrifice fly in which the batter hits the ball deep enough that while the out is recorded, a runner has time to tag up and advance a base. Since making an out is an output, but an undesirable output, strikeouts and fly outs are included in the model as inputs.

Table 1: Inputs and Outputs to the DEA Analysis

Inputs	**Outputs**
P	1B
K	2B
F	HR
	BB

Returning to the original plate appearances. An adjustment is necessary since the number of plate appearances also includes the strikeouts and the fly outs. To avoid the possibility of multicollinearity, plate appearances are reduced by the strike outs and fly outs. Thus $\mathrm{P} = \mathrm{PA} - \mathrm{K} - \mathrm{F}$.

Assuming that there are N ball players to be compared, the primal LP model for this DEA for player k is

$$
\begin{aligned}
&\max z = \nu_1 1\mathrm{B}_k + \nu_2 2\mathrm{B}_k + \nu_3 \mathrm{HR}_k + \nu_4 \mathrm{BB}_k - \epsilon\left(\mu_1 \mathrm{P}_k - \mu_2 \mathrm{K}_k - \mu_3 \mathrm{F}_k\right) \\
&\text{subject to} \\
&\mu_1 \mathrm{P}_k + \mu_2 \mathrm{K}_k + \mu_3 \mathrm{F}_k = 1 \\
&\nu_1 1\mathrm{B}_j + \nu_2 2\mathrm{B}_j + \nu_3 \mathrm{HR}_j + \nu_4 \mathrm{BB}_j - \mu_1 \mathrm{P}_j - \mu_2 \mathrm{K}_j - \mu_3 \mathrm{F}_j \leq 0 \qquad j = 1, 2, \ldots, N \qquad (11)
\end{aligned}
$$

A separate linear program would be created and solved for each player.

As mentioned previously, the analysis requires not only the solution of the primal model but the dual model as well. The primal will provide the relative efficiency of each player, and the amount that each output would need to be increased in order for an inefficiency player to reach efficiency. But increasing an output may not be a simple task. A manager would be thrilled to tell a player that they must simply hit more home runs to be more efficient. But can that actually be done? This is where the dual model enters.

The dual can be created by swapping inputs and outputs for each DMU. Since the original goal is to maximize relative efficiency and thus determine the minimum possible values for the inputs, for the dual the goal is minimization in order to determine, using the slack values, what the maximum of an input could be. The difference then provides the amount an input would need to be decreased.

From a practical perspective this may be more implementable. If a player is inefficient because of the number of strike outs then the coaching staff might be able to work with them on techniques to reduce strike outs. Similarly, if they are inefficient because of the number of fly outs, they could perhaps work on keeping the ball down, or not swinging for the fence as often.

The dual model for the batting DEA for player k is

$$\min z = \sum_{j=1}^{N} \Theta_j$$

subject to

$$\sum_{i=1}^{N} 1\mathrm{B}_i\Theta_i \geq 1\mathrm{B}_k \qquad \sum_{i=1}^{N} 2\mathrm{B}_i\Theta_i \geq 2\mathrm{B}_k \qquad \sum_{i=1}^{N} \mathrm{HR}_i\Theta_i \geq \mathrm{HR}_k \qquad \sum_{i=1}^{N} \mathrm{BB}_i\Theta_i \geq \mathrm{BB}_k$$

$$\sum_{i=1}^{N} \mathrm{P}_j\Theta_i > 0 \qquad \sum_{i=1}^{N} \mathrm{K}_j\Theta_i > 0 \qquad \sum_{i=1}^{N} \mathrm{F}_j\Theta_i > 0 \tag{12}$$

Results

Data was collected for the 2022 season from the website *Baseball Savant*(*MLB Batting Statistics*, 2022). This included all batting data on 644 players. Of these, the model was run on the 276 players who had at least 300 plate appearances during the 2022 season. This required running 276 separate models - one for each player.

There were 53 of the 276 players who received a relative efficiency of greater than 0.999 or 99.9% . The top ten are shown in 2.

Table 2: Top Ten Players in Relative Efficiency

Player	Net PA	K	F	1B	2B	HR	BB	WAR
Jean Segura	294	58	35	79	9	10	25	1.8
Tim Anderson	257	55	39	81	13	6	14	1.3
Aaron Judge	445	175	76	87	28	62	111	10.6
Michael Harris	302	107	32	74	27	19	21	5.3
Yordan Alvarez	386	106	69	76	29	37	78	6.8
Mookie Betts	412	104	123	76	40	35	55	6.4
Vladimir Guerrero	527	116	63	108	35	32	58	3.9
Nolan Arenado	450	72	98	90	42	30	52	7.9
Teoscar Hernandez	328	152	55	72	35	25	34	2.8
William Contreras	241	104	31	58	14	20	39	2.7

The data for each is shown including an additional statistic, *WAR*. A statistic that has become popular over the last few seasons, WAR is an acronym for Wins Above Replacement. It is considered a measurement of the value of the player in winning games. The theory is that WAR measures how many more wins a team would have with that particular player instead of a replacement at the same position. It is a measurement of hitting, running, and fielding. It is included in these results as a means of comparing the efficiency of the player with a more common statistic.

In addition to the top ten most efficient players, it might be interesting to identify the bottom then. They are shown in table 3. Since these players are identified as inefficient batters, the results also show the number of additional hits or fewer outs each would need to become efficient.

These results show a higher WAR for efficient players as compared to the inefficient players, but this is expected. The results of the inefficient players do not show large slack values in the primary outputs - hits, but they do show significant surpluses in the inputs. Recall that strikeouts and fly outs while actually outputs of hitting are undesirable and were thus analyzed as inputs. An interpretation of this may be that the ratios of types of hits are consistent with more efficient players, but that these batters required more plate appearances and thus struck out more often, or popped out to the outfield. This provides a means of identifying where the inefficiency is occurring.

Table 3: Bottom Ten Players in Relative Efficiency

Player	Efficiency	Net PA	K	F	1B	2B	HR	BB	WAR
Austin Hedges	64.7	215	78	45	37	4	7	25	−.4
		(−76.0)	(−38.3)	(−15.9)	(0.0)	(5.0)	(0.3)	(0.0)	
Abraham Toro	68.1	212	65	75	36	10	7	22	−.8
		(−67.7)	(−20.8)	(−40.4)	(0.0)	(0.0)	(0.0)	(0.0)	
Jonathan Schoop	69.6	343	107	60	62	23	11	19	0.1
		(−104.4)	(−32.6)	(−18.3)	(0.0)	(0.0)	(0.0)	(0.7)	
Martin Maldonado	71.7	214	116	49	37	12	15	22	0.3
		(−60.5)	(−48.9)	(−49.8)	(0.0)	(0.0)	(0.0)	(3.5)	
Rougned Odor	72.3	288	109	75	53	19	13	32	−.4
		(−79.8)	(−39.5)	(−27.4)	(0.0)	(0.0)	(0.0)	(3.5)	
Nick Allen	72.4	225	64	37	45	13	4	19	1.0
		(−62.1)	(−25.1)	(−10.2)	(0.0)	(0.0)	(0.7)	(0.0)	
Victor Caratini	72.4	216	67	31	33	12	9	31	0.5
		(−59.6)	(−18.5)	(−8.6)	(0.0)	(0.0)	(0.0)	(0.0)	
Jesus Sanchez	73.7	210	92	41	37	14	13	26	1.0
		(−55.3)	(−34.8)	(−10.8)	(0.0)	(0.0)	(0.0)	(0.0)	
Jeimer Candelario	74.3	298	109	60	59	19	13	28	0.6
		(−76.7)	(−29.3)	(−15.4)	(0.0)	(0.0)	(0.0)	(0.0)	
Tony Kemp	75.9	400	69	89	84	24	7	45	1.1
		(−96.2)	(−16.6)	(−21.4)	(0.0)	(0.0)	(1.2)	(0.0)	

Another action can be taken using these results. While hitting a baseball is considered one of the most difficult feats in sport, not hitting it is far easier. If a batter is inefficient because of a low number of bases on balls it may be that they swinging away instead of taking a pitch. There are numerous pros and cons on taking a first pitch, but one is that working the count increases the number of walks. Of course it needs to be said that it may also increase the number of caught looking strikeouts.

The DEA analysis can identify the players who are not walking enough. The top five are shown in table 4.

Table 4: Players with the largest difference between expected and actual number of bases on balls

Player	Efficiency	Net PA	K	F	1B	2B	HR	BB	WAR
Luis Rengifo	88.1	362	79	70	86	22	17	17	2.4
		(−43.0)	(−9.4)	(−9.3)	(0.0)	(0.0)	(0.0)	(24.4)	
Salvador Perez	95.6	280	109	84	66	23	23	18	2.7
		(−12.3)	(−4.8)	(−20.4)	(0.0)	(0.0)	(0.0)	(19.3)	
Jeremy Pena	90.3	358	135	65	88	20	22	22	4.8
		(−34.7)	(−13.1)	(−10.6)	(0.0)	(1.3)	(0.0)	(18.5)	
Pete Alonso	91.6	458	128	99	95	27	40	67	4.4
		(−38.6)	(−10.8)	(−17.5)	(0.0)	(0.0)	(0.0)	(13.3)	
Bobby Witt	93.1	376	135	121	93	31	20	30	0.8
		(−5.8)	(−9.3)	(−41.4)	(0.0)	(0.0)	(0.0)	(13.1)	

Conclusions

Analytics in baseball has been done for over one hundred and fifty years. But the success of implementing analytics as a means of decision making by Billy Beane and the Oakland A's has completely changed the game. From pitch counts

to balls and strikes, every aspect of the game is being analyzed. Most of the statistical calculations use variations of the tradition descriptive statistics. But data envelopment analysis can provide an alternative approach to analyzing America's Pastime.

DEA implements linear programming to determine the relative efficiency of a set of decision making units. In baseball the DMU can the individual batter and the efficiency in how they convert inputs - plate appearances - to outputs - singles, doubles, home runs, and bases on balls. By calculating slack and surplus values for each variable, it is then possible to determine where a batter is inefficient.

References

Anderson, T. R., & Sharp, G. P. (1997). A new measure of baseball hitters using DEA. *Annals of Operations Research*, *73*, 141-155.

Chadwick, H. (1872). *De Witt's base-ball guide for 1872*. New York: Robert M. de Witt.

Cook, E. (1964). *Percentage baseball*. Cambridge, MA: MIT Press.

Jabhanshahloo, G. R., Lotfi, F. H., Shoja, N., Razavyan, S., & Tohidi, G. (2005, October). Undesirable inputs and outputs in dea models. *Applied Mathematics and Computation*, *169*, 917-925.

James, B. (1982). *Baseball abstract*. Morristown, NJ: Ballantine.

Lewis, M. (2003). *Moneyball: The art of winning an unfair game*. New York: W.W. Norton & Company.

Mazur, M. J. (1994). Data envelopment analysis: Theory, methodology, and application. In A. Charnes, W. W. Cooper, A. Y. Lewing, & L. M. Seiford (Eds.), (p. 369-392). New York: Springer Science and Business Media.

Ramanathan, R. (2003). *An introduction to data envelopment analysis: A tool for performance measurement*. New Delhi: Sage Publications.

MLB Batting Statistics. (2022). Retrieved from `https://baseballsavant.mlb.com`

Wulf, S. (2012, September). *Davey Johnson: Nationals' pleasure.* Retrieved December 25, 2022, from `https://www.espn.com/mlb/story/_/id/8383480`

Wyner, A. (2019, February 21). Changing the game: How data analytics is upending baseball. *Knowledge at Wharton*. Retrieved from `https://knowledge.wharton.upenn.edu/article/analytics-in-baseball/`

Past, Present, and Future of Digital Transformation-A Literature Review

Abstract

This study explores the development track of the "Digital Transformation" (DT) literature, searches the academic database (Scopus) and related academic journals to explore the context and trends of DT.

The effective composition analysis condition is 3,636 journal documents obtained from the academic database, and a citation network is established for the citation data of the digitally transformed academic documents. Then, the main path analysis is adopted for the citation network so that the complex network journal articles can form a systematic main path.

In addition, the abstracts of all relevant papers in the literature are analyzed by cluster analysis and word cloud analysis to determine the keywords with high frequency to understand the research topics that the related researchers are more interested in, and the final results obtain the four main research topics. They are technological innovation and performance, Industry 4.0, medical education process, and information and communication technology. It also counts the growth trajectory of the number of people who pay attention to this topic and the number of articles every year and reveals which important journals related papers are published in.

This study compiles the literature related to digital transformation over the years, demonstrates an integrated analysis step, provides a unique method to describe the application and development of the theory and practice of digital transformation, and summarizes the arguments of researchers in each period to keep pace with the times as a reference for future researchers interested in this topic.

Keywords: ***digital transformation, digitization, main path analysis, word cloud analysis, development trajectory***

1. Introduction

In recent years, people from all walks of life have paid close attention to the topic of digital transformation (Digital Transformation; DT), which has quickly aroused a trend in academia and has great sensation and discussion space in the industry, especially in the world suffering from the ravages of COVID-19 and the after the impact, the global development and implementation of digital transformation has been accelerated (Candelo et al., 2021). There is growing evidence that digital transformation has the potential to improve operations and stimulate corporate growth (Liu & Liu, 2021; Guo & Xu, 2021), has become an effective way for companies to remain competitive (Yang et al., 2021), and has the characteristics of ease of use (Buck et al., 2021). As a result, various industries and enterprises are irresistible to the issue of digital transformation. Teece, 2007) to ensure that the organization can obtain the greatest benefit from this issue. In the past two years, scholars' research on digital transformation has increased, and it has been shown that the issue of digital transformation has deeply affected the formulation and operation of corporate strategies (Liu & Lu, 2021; Hess et al., 2016). Many scholars have also clearly shown that digital transformation is positively correlated with business performance (Guo & Xu, 2021; Tsuu & Chen, 2021). In addition, there are also studies showing that the introduction of digital transformation can also increase customer experience and simplify operational processes. and even create new business models for organizations (Climent & Haftor, 2021; Warner & Wäger, 2019).

Although these studies have contributed to our understanding of digital transformation in specific fields, we currently lack a multifaceted understanding of this phenomenon and systematic impact analysis research (Vial, 2019), so there is still a lack of promotion of digital transformation. Empirical research on digital transformation (Liu & Liu, 2021) shows that most companies want to pursue DT but are often unable to do so. According to statistics, most business leaders agree that "DT is necessary to keep up with business." The economy is changing", but only 23% of business leaders have a complete DT strategy, showing that companies have a closeness to DT. Therefore, this study summarizes the 3,636 journal articles in the database and uses the main path analysis, cluster analysis and word cloud analysis methods to systematically summarize, analyze and study the related diversity research on DT. The essence of several questions, such as: What is DT? What are the factors? What is the transformation process? What are the barriers to DT? Are there any similarities among these related studies? What fields does each study apply to? What is the future trend of DT? In addition, do researchers focus on the same field or do research across fields and levels.

With the continuous development of DT in various fields and levels, the importance of DT to enterprises cannot be overemphasized. The concept of DT has also emerged in recent years. The essence still makes people feel like a hundred flowers contend and disagree, so it is increasingly important to deeply explore the consistency and integration of these different research fields. Therefore, this study believes that it is necessary to systematically clarify or

explain the above problems. This research hopes to provide a reference for subsequent researchers or enterprises after in-depth exploration of the overall development trajectory of DT and the relationship between theoretical concepts.

2. Conceptual Background

Soucek (1968) first mentioned the term DT in the article "Million channel pulse height analyzer through pseudorandom" published in Nuclear Instruments and Methods. However, at that time, because it only described the transformation of data in the nuclear spectroscopy laboratory, the term DT did not receive much attention at that time. In recent years, an increasing number of academic fields have continued to explore DT. Why? How does a company build DT? How does DT affect a company's business strategy and successfully change its business model or business performance?

Research on DT covers a wide range of related fields; for example, business management scholars study how enterprises use digital technology to create and distribute business value (Climent & Haftor, 2021), and human resources scholars study how human resources mediate the relationship between corporate strategy and DT. relationship (Nicolás-Agustín et al., 2021), and educators have explored the extent to which DT is applied in university higher education (Mikheev et al., 2021; Safiullin & Akhmetshin, 2019; Bond et al., 2018). In terms of sociologists, scholars have also found that DT not only affects the national economy but also changes social relations (Okhrimenko et al., 2019), and public health scholars have also studied the role of DT in the health care system (Agarwal et al. al., 2010; Gray et al., 2015). In the information section, scholars have studied the impact and role of IT departments and their services in DT (Du et al., 2016; Hansen et al., 2011), while organizational scholars have also incorporated DT into organizational culture and organizational change (Hanelt et al., 2021), and economists have focused on how DT affects the performance of the circular economy (Tsou & Chen, 2021; Bag et al., 2020).

Through a review of 282 papers, Vial (2019) concluded a DT framework consisting of eight influencing factors. Tijan et al. (2021) conducted a 139 literature review describing barriers to DT. Rêgo et al. (2021), through a systematic review of the DT literature, indicated that research on the relationship between DT and strategic management is still in its infancy and divided the literature into external environment, internal environment, strategy formulation, strategy implementation, and evaluation. and control and feedback and learning in six main research areas, Majchrzak, Markus & Wareham (2016) discuss the impact of ICT on society through a study of 14 papers in the journal, Matarazzo et al. (2021) study SMEs The impact of DT on customer value creation, research shows that DT is conducive to business model innovation. The abovementioned literature review aims to unify the limited literature on DT in a specific field. Although there is more understanding of this topic, it is impossible to understand the entire development of DT. Therefore, this study will conduct a retrospective

analysis of the relevant literature on DT from 1968 to November 2021 on WOS-Scopus, which can explore the overall development trajectory and main research topics of DT and analyze the relationship between academia and industry. Knowledge Diffusion for DT.

3. Review Methodology

This study combines main path analysis , cluster analysis and text mining tools to completely analyze the development trajectory of DT. The main path analysis is to systematically organize the citation relationships between past academic documents to explore the overall development trajectory of DT. Cluster analysis is the main research topic used to find clusters of literature related to DT. Mining refers to the statistics of the frequency of occurrence of words and nouns in the abstracts of DT literature and their key words, mainly to determine the research focus of researchers.

3.1 main path analysis

Main path analysis is a citation analysis of documents that determines the citation relationship between documents. The main path concept was first introduced by Hummon & Doreian (1989). The basic concept of main path analysis is to link citation data and social network analysis to determine the development trajectory of related knowledge diffusion to simplify the complex citation network among papers in the database into a main path. To illustrate the main theoretical development and application process in this field. Main path analysis has a wide range of uses. Hummon has performed a large number of citation analysis studies with many scholars, such as Hummon, Doreian & Freeman's (1990) productivity literature citation analysis, Hummon & Carley's (1993) social network field citation analysis and Carley, Hummon & Harty's (1993) citation analysis of conflict resolution. To solve this problem, Batagelj & Mrvar (1998) developed Pajek software, which graphically presents complex citation main paths. However, traditional main path analysis still has some limitations, and some important links may be missed. Therefore, Liu & Lu (2012) proposed a key-route main path (Key-Route Main Paths) analysis method, which is to determine first for the connection with the largest weight, find the main path forward from the front end of the connection, and then find the main path from the back end of the connection. Extending the main path, the main advantage of this method can be seen in the phenomenon of separation and integration of the overall development of theory or technology and can clearly describe the process of knowledge development at different stages.

The purpose of this study is to deeply explore the nature of DT. Therefore, to understand the development trajectory of knowledge diffusion related to "DT", it is more appropriate to use main path analysis as a research method to observe this issue.

3.2 Cluster analysis

Cluster analysis is used to find groups with certain correlation characteristics from a large

amount of original data. It is a multivariate analysis program. There is a large difference in this characteristic of the cluster analysis. In short, the phenomenon of the characteristic variability of the cluster analysis is the similarity within the group and the large degree of dissimilarity between the groups. If the researcher wants to observe whether the quality of the data under study is similar, cluster analysis is a very useful technique in the data exploration and analysis part. Due to the different analysis methods used in cluster analysis, the results will be different. When different researchers perform cluster analysis on the same observation, the number of clusters determined may not be the same. Therefore, the use of cluster analysis is usually more inclined to use research in exploratory analysis. The application of cluster analysis of observations is to divide independent observations into different groups (groups), and when using cluster analysis, the characteristics of the groups to which the observations belong are not known at the moment. The observer does not know how many clusters the independent observations can be divided into. If the number of clusters is unknown, the characteristics of the clusters are naturally unknown. The grouping logic of cluster analysis is to use "numerical taxonomy", and the classification criterion is not determined by the researcher in advance but to classify the measured samples, and after classification, the samples in the same group have a high degree of classification. The similarity and homogeneity of the samples are high, while the samples in different groups have high heterogeneity; the focus of this method is not to estimate the variation. Instead, the observations are grouped by using the differences in the variance of the variables.

This research mainly reviews the relevant literature on "DT" in recent years and determines the essence of the issue. Therefore, group analysis is used to group the obtained literature on DT over the years, and many groups are the main literature in this field. It is expected to determine the differences in variables from each group; therefore, this study chooses group analysis as the research method.

3.3 Text Mining

Text mining is from unstructured textual data. Among them, the meaningful information hidden behind is discovered, which can be used as a reference for decision-making. Although the text in the text may be ambiguous or the text may be different from the emotions hidden behind it, text mining has the ability to systematically identify, capture, manage, integrate and apply the knowledge hidden behind text data, so the use of text mining can still obtain the sentiment and opinion of the article. Then, the VOSviewer system program was used to observe the relationship between the representative characters to confirm the relationship between them. This research is based on a statistical analysis of the literature related to "DT" over the years. The purpose is to determine the research focus of the researchers. Therefore, the text mining of the abstracts and key words of the literature in each group is aimed at determining the attributes of the researcher's key issues from the unstructured text data in the relevant literature of each group.

3.4 Database analysis

The basic data source is extracted from the database Scopus. The search scope was related journal literature published from 1968 to 2021 (end of November). There are 3,636 valid data points. This article serves as a dataset for follow-up research. According to the statistical results, in the past three years, approximately 80% of the papers have been published on this topic. It can be seen that the DT topic has received attention in the past three years. Especially in the past two years, more than a thousand papers have been published every year. Figure 1 shows that the DT topic is an emerging topic and is full of vitality, and the number of papers published from 2019 to 2021 (as of November) has a clear growth trend, which further shows that this topic is attracting attention from scholars, as shown in Figure 2.

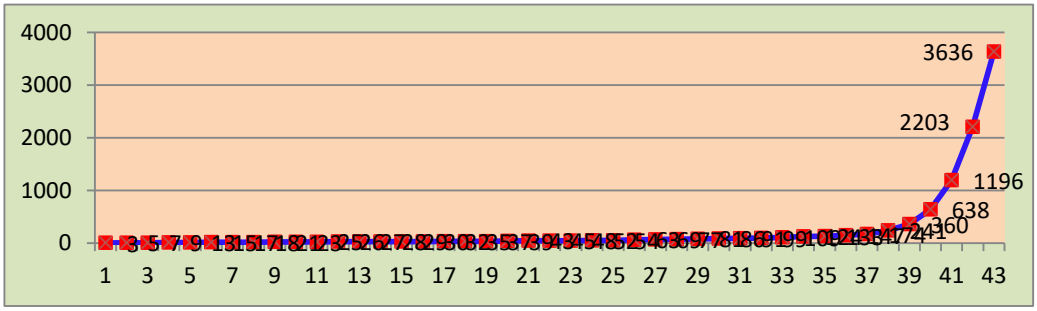


Figure 1 Number of publications related to "DT" over the years

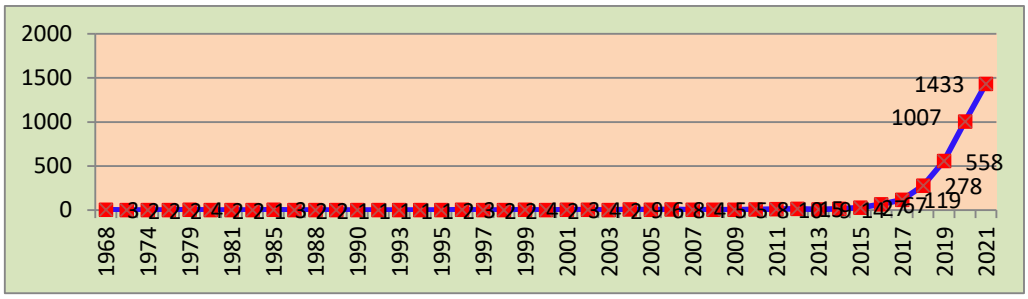


Figure 2 Number of publications related to "DT" over the years

Table 1 below uses the g-index as the basis for ranking the top 20 journals with the most influence on the topic of "DT", as well as the cumulative number of related papers published, the number of papers published after 2000, the journals published, etc. Statistics. According to the statistics in the table below, the top five major journals that publish articles on DT are sustainability (Switzerland), Journal of Business Research, Technical Forecasting and Social Change, Idea Access, and Mis Quarterly Executive. We find that each of the top 20 journals was published after 2000, the first-ranked journal is sustainability (Switzerland), and the articles in this journal are among all the current journals in terms of the number of articles published or active years. It is in the leading stage, and although the h-index of this journal is not much different from other journals, it is observed that its g-index is relatively high, indicating that papers with high citations are published in this journal, indicating that the journal is affected by

DT. Scholars in the field are concerned, and it can be described as the mainstream journal in this field. While technical forecasting and social change and ieee access have more publications than the journal of business research, their h-index and g-index are not relatively outstanding, indicating that the journal's articles are cited relatively more than the journal of business research. There are few business research journals. In addition, we can also see from the top 20 journals that the sixth place mis quarterly executiveiv and the eighth place international journal of information management are exposed to DT issues earlier than other journals; it is worth mentioning that as mentioned above, DT belongs to It is an emerging issue and is still in the exploratory stage. Therefore, the publication of academic articles related to DT issues in the future may also affect the current ranking.

Table 1 Top 20 influential journals

Total papers	Papers after 2000	g-index	h-index	Years	Journals
142	142	28	18	2018~2021	sustainability (switzerland)
35	35	23	11	2019~2022	journal of business research
42	42	23	10	2018~2022	technological forecasting and social change
38	38	18	9	2017~2021	ieee access
17	17	17	11	2011~2021	misquarterly executive
15	15	15	9	2016~2021	business horizons
15	15	15	9	2011~2021	international journal of information management
15	15	14	6	2018~2021	business process management journal
14	14	14	5	2020~2021	journal of manufacturing technology management
13	13	13	7	2019~2021	industrial marketing management
12	12	12	6	2016~2021	ieee engineering management review
12	12	12	4	2017~2021	international journal of innovation management
20	20	12	6	2018~2021	journal of medical internet research
11	11	11	8	2018~2021	california management review
11	11	11	8	2019~2021	government information quarterly
11	11	11	6	2017~2021	production planning and control

10	10	10	4	2019~2022	advanced engineering informatics
28	28	10	6	2019~2021	applied sciences (switzerland)
13	13	10	5	2017~2021	journal of business strategy
10	10	10	5	2020~2021	technologynovation

Note: h-index means that h articles have been cited at least h times; g-index is mainly to improve the shortcomings that h-index cannot fully reflect the highly cited literature, that is, all papers published by an author are ranked from high to low according to the number of citations. When the cumulative number of citations of the first g papers is greater than or equal to g2, it is the g-index of the journal or author.

4. Performance Analysis

This research uses key-route main path analysis to deeply explore the overall development trajectory of DT and, through the linking trajectory of the citation relationship between papers, to understand what research the academic community has mainly engaged in in this field in the past and present. The main path of the key route is shown in Figure 3. We found that influential papers in this field can be divided into two parts, one of which is the construction of DT. In relation to performance, the other (left in the picture below) explores what DT is. The green nodes in the figure represent the source points, and the blue nodes are the sink points. The thicker the connecting line is, the greater the influence. In other words, the source point is the starting point, and the sink point is the ending point. The code name of each node is the surname of the first author, plus the first letter of the surname of other coauthors, and the number represents the publication year of the paper.

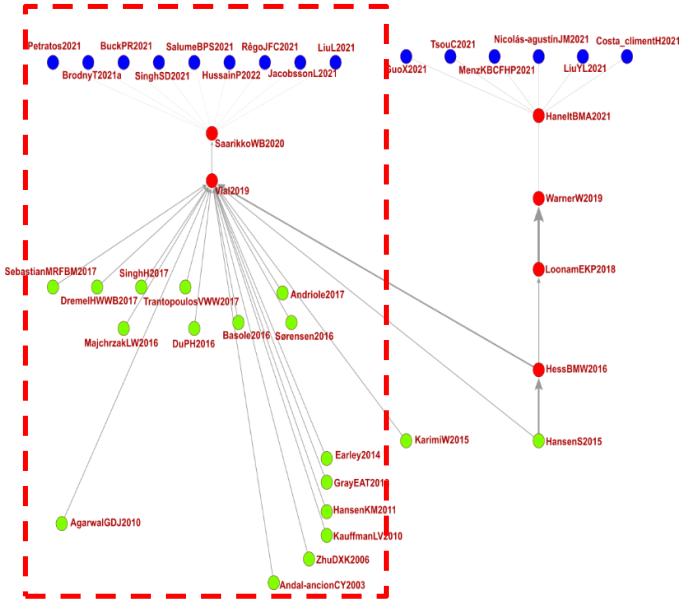


Figure 3 The overall development trajectory of academic literature

On the right side of Figure 3 are related papers on constructing DT and corporate performance. Beginning in 2015, scholars at this time emphasized that enterprises use network technology to combine online and offline omni-channel retail channels to provide customers with a seamless experience (Hansen & Sia, 2015). Research on related topics, including how corporate executives face DT (Hess et al., 2016), the actions required by organizational management to implement DT (Loonam et al., 2018), how companies structure the dynamics of DT capabilities (Warner & Wäger, 2019), driving DT from the perspective of organizational change (Hanelt et al., 2021), the relationship between DT and business models (Climent & Haftor, 2021), the ability of industries and enterprises to adapt to DT (Liu, Yang, Liu, 2021), the relationship between DT and corporate strategy (Menz et al., 2021), the corporate human resources required for DT (Nicolás-Agustín et al., 2021), the impact of DT and organizational innovation on companies' performance impact (Tsou & Chen, 2021) and the impact of DT on corporate operations and financial performance (Guo & Xu, 2021), shows that DT has always been closely related to corporate performance.

In the exploration of DT, in recent years, relevant scholars have tried to explore the impact and application of DT from various aspects and perspectives, such as health care (Agarwal et al., 2010), information security (Lowry et al., 2021; Chen et al., 2021), SMEs (Brodny & Tutak, 2021; Candelo, Casalegno, & Civera, C., 2021), enterprise dynamic capabilities (Karimi, & Walter, 2015), technological innovation (Zhu et al., 2006), IT governance (Joshi et al., 2021), organizational change (Hanelt et al., 2021), and business strategy (Liu, & Lu, 2021; Menz et al., 2021). From the above, it can be clearly observed that the DT issue is still spreading in all walks of life.

The overall development trajectory of academic literature is to use the main path analysis to determine the most influential papers from the past literature from the macroscopic perspective and to form the development trajectory through the path of citation relationships. From Figure 3, we can easily find that DT has a high impact on business operations. To further understand the situation of DT in other applications, we use clustering technology to explore other application areas from a more nuanced perspective.

From the overall development trajectory of academic literature, it can be seen that the most influential aspects of DT belong to the research on the construction and performance of DT, but we are also eager to understand the development of DT in other subfields, so we put all the literature further. Do grouping, select representative literature from groups with more literature, take their abstracts and key words for word cloud analysis, and provide a name for the group's research topic.

Table 2 Research topics in academic literature

Group	Research Theme	Number Of Articles	Keyword/Word	Word Cloud Analysis
Cluster 0	Technological Innovation And Performance	244	Performs (24) Technology (22) Model (13) Innovation (13) Develop (12)	
Cluster 1	Industry 4.0	63	Industry 4.0 (203) Technology (128) Manufacture (97) Model (71) Process (68)	
Cluster 2	Medical Education Process	28	Process (10) Medical (8) Covid-19 (8) Education (7) Social (7)	
Cluster 3	Newsletter Technology	33	Technology (14) Environmental (12) Literature (11) Ict (11) Impacts (9)	

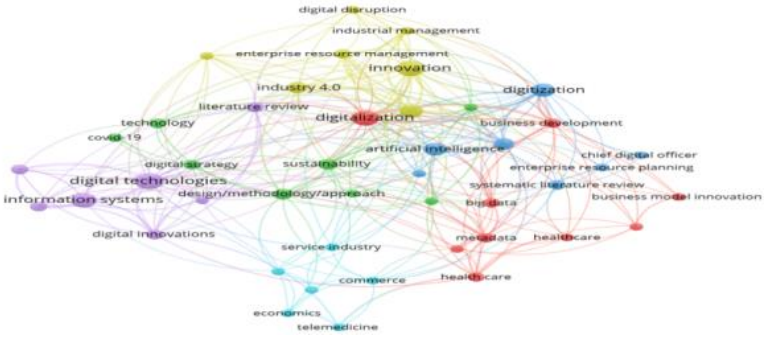
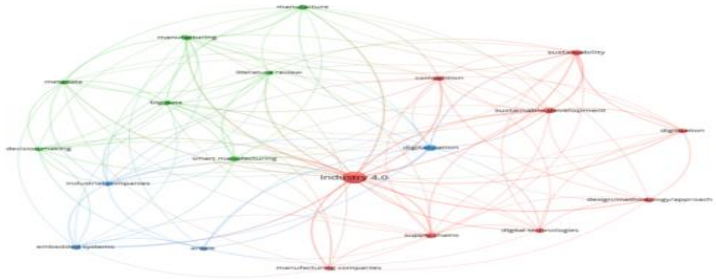
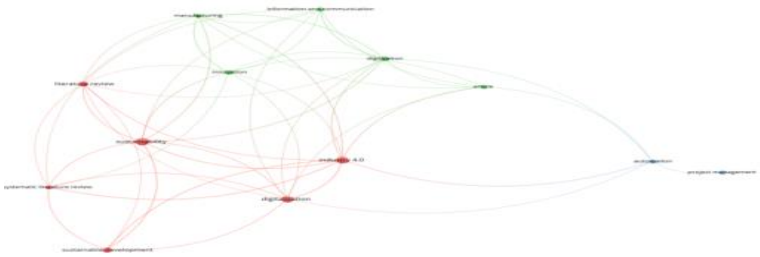
Note: The numbers in parentheses are the occurrences of keywords/words in the group papers

Through the above steps, a total of 4 groups were obtained by cluster analysis. Table 2 lists the research topics, number of papers, keywords/words, and word cloud analysis results of each group. The word cloud is a statistical analysis of all the paper abstracts and keywords/words in the group and is then presented in the form of a word cloud. Among the 4 major groups, which are closely related to business operations, they are technological innovation and performance and Industry 4.0; the other 2 groups are medical education process and information and communication technology. Innovation and performance are most popular with DT scholars.

We then use VOS viewer to observe the relationship diagram of each group of words (Table 3). The first group is related to activity. From this group, we know that the hot area (purple band) is innovation and performance, and we can also observe digital information supporting this area. The second group extends outward with Industry 4.0 as the center point,

the third group is purely related and discusses the relationship between university education and medical care, and the fourth group is information systems that drive DT. related fields.

Table 3 The relationship between the first four groups of word clouds

research theme	Word Cloud Relationship Analysis
Technological performances	
Industry 4.0	
medical education process	
Newsletter technology	

After we understand the research topics of these 4 major groups, we want to further understand the connection between the 4 major groups and the overall main path. Therefore,

we use the branch function so that only the 4 major groups are marked with 4 in the overall main path.The results of the papers linked by the large group are shown in Figure 4; the application of digital technology and innovation has been described in the previous section, and the other two groups related to Industry 4.0, medical education process and information and communication technology will be described later.

From the diagram of the relationship between the subfield and the overall main path, it can be seen that the four groups of DT are involved with each other, which shows that there is a certain relationship between the groups. Among the issues, there is a relationship between the groups.

First, after discussing "DT and Innovation", we then discuss the "Industry 4.0" group, which was formed in 2019, and at this time began to describe the components of Industry 4.0, as well as Industry 4.0 and DTal relationships (Castelo-Branco et al.,2019; Frank et al.2019), which subsequently evolved to describe how DT is applied in Industry 4.0 (Culot, 2020). Follow-up researchers follow the accumulation of their knowledge, and their research includes directly pointing out the functions of DT in the framework of Industry 4.0 (Calabrese et al., 2020). Based on the main path analysis, it can be obtained that the main path is Culot (2020) as the main core; therefore, we know that this group of DT focuses on the cooperation and development of applications for Industry 4.0.

The third group is named "Medical Education Process". This group originated in 2018. Maltese (2018) described the university's promotion of digital services to the public, but the inherent data fragmentation and data diversity increased the promotion of digital services. difficulty, and Ramírez-Montoya (2020), when researching university education and COVID-19, found that DT (virtualization, training, infrastructure, connectivity, culture, management, open education) and educational innovation (new processes, products, services, knowledge and research) can address this challenge. Safiullin & Akhmetshin (2019) also show that the digital economy has opened up new advantages and disadvantages for universities and increased their competitiveness. We know that this group mainly starts from the perspective of universities and extends from digitalization to the application of infectious diseases.

Another group originated in 2018. Fritzsche et al. (2018) studied the impact of climate change and energy on Industry 4.0. Kunkel & Matthess (2020) described the importance of information and communication technology (ICT), noting the impact of ICT on digital and industrial policy sustainability as predicted. They also established an ICT framework that distinguishes ICT from indirect environmental impacts arising from the application of ICT in various production processes and economic activities. We can clearly see that the core of this group is "ICT", focusing on system models and methods, so we named this group "Information and Communication Technology".

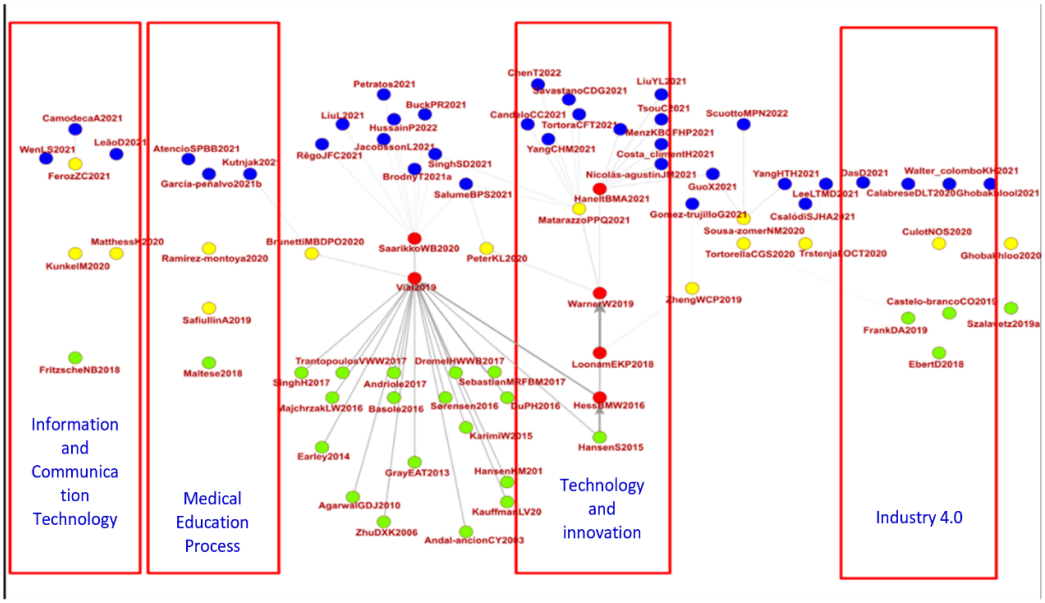
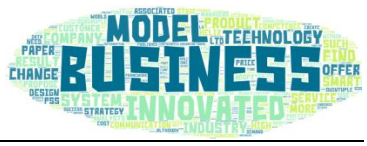


Figure 4: The diagram of the relationship between the subfield and the overall main path

To have a more complete understanding of the current overall situation of DT, we also used word cloud text to analyze the literature of the remaining 16 groups, trying to explore what are the research themes of these 16 groups? The results are BIM, SME research, Internet of Things, transformation performance, sales value, system integration, higher education, digital innovation, supply chain management, knowledge reorganization, risk assessment, production process management, AI artificial intelligence, business model innovation, big data and digital transformation strategies, etc. 16 groups, as shown in Table 3. We can see from the table below that the application of DT is quite extensive, but the number of published papers is relatively small. Although these papers have their contribution and value in this specific field, this research is based on the influence of the literature. Gao's main path analysis reflects the overall development trajectory of academic literature, so the 16 groups are not included.

Table 4 Research Topics of Academic Research Clusters 5 to 20

research topic	Number of papers	Word Cloud Analysis
Building Information Model	42	

Small and Medium Enterprise Research	33	
Internet of Things	20	
Transformational performance	27	
sales value	13	
System integration	22	
higher education	13	
digital innovation	20	
supply chain management	12	
knowledge reorganization	11	
risk assessment	17	
Production process management	12	
AI artificial intelligence	12	
business model innovation	14	

big data	13	
DT Strategy	12	

5. Conclusion

This research combines the three methods of main path analysis, cluster analysis and text mining to deeply explore the overall development trajectory of the DT literature. The main path analysis results show that the application of DT is still mainly based on performance. In recent years, an increasing number of scholars have rapidly invested in research in the field of DT, which has made the application of DT increasingly extensive. In addition to the corporate world, its application scope has even extended to the education industry and the construction industry. Cluster analysis was combined with word cloud analysis to identify four major themes of DT, namely, technological innovation and performance, Industry 4.0, medical education process, and information and communication technology.

Although the application of DT research is multilevel and has developed rapidly in the past few years due to attracting the attention of many scholars, there is no relevant research that systematically summarizes the application of DT. Therefore, we made several contributions. First, we reviewed all the literature on DT and confirmed the overall development trajectory and application of DT through main path analysis, cluster analysis and text mining. Second, we also use these three types of analysis to confirm the relationship between the subdomains of DT and the overall main path. We organize these literatures through such a comprehensive framework. We use these studies to review the application of multilevel analysis of DT. empirical research. Third, we sort out the context of DT from the subgroups of DT and determine the coherence and conflict points of each subgroup in the empirical results. Fourth, we identify emerging research themes through main path analysis to highlight areas for future research. Finally, we hope that through this research, subsequent researchers can deepen their understanding of DT and expand the application of DT, which can be used as a reference for formulating future enterprise-related strategies.

References

Agarwal, R., Gao, G., DesRoches, C., & Jha, A. K. (2010).” Research commentary—The digital transformation of health care: Current status and the road ahead”. Information systems research, 21(4), 796-809.

B., S., V., B., & K., C. (1968) “Million Channel Pulse Height Analyzer through Pseudo-Random Digital Transformation”. NUCLEAR INSTRUMENTS AND METHODS, (66), 202~212.

Bag, S., Wood, L. C., Mangla, S. K., & Luthra, S. (2020). “Procurement 4.0 and its implications on business process performance in a circular economy”. Resources, conservation and recycling, 152, 104502.

Bond, M., Marín, V. I., Dolch, C., Bedenlier, S., & Zawacki-Richter, O. (2018). “Digital transformation in German higher education: student and teacher perceptions and usage of digital media”. International Journal of Educational Technology in Higher Education, 15(1), 1-20.

Buck, C., Probst Marques, C., & Rosemann, M. (2021).” Eight Building Blocks for Managing Digital Transformation”. International Journal of Innovation and Technology Management, 18(05), 2150023.

Batagelj, V., & Mrvar, A. (1998). “Pajek-program for large network analysis”. Connections, 21(2), 47-57.

Climent, R. C., & Haftor, D. M. (2021).”Business Model Theory-Based Prediction of Digital Technology Use: An Empirical Assessment”, Technological Forecasting and Social Change, 173, 1~14.

Carley, K. M., Hummon, N. P., & Harty, M. (1993). “Scientific influence: An analysis of the main path structure in the Journal of Conflict Resolution”. Knowledge, 14(4), 417-447.

Candelo, E., Casalegno, C. G., & Civera, C. (2021). “Digital transformation or analogic relationships? A dilemma for small retailer entrepreneurs and its resolution”. Journal of Strategy and Management.

Chen, W., Zhang, L., Jiang, P., Meng, F., & Sun, Q. (2021). “Can digital transformation improve the information environment of the capital market?” Evidence from the analysts' prediction behavior. Accounting & Finance.

Castelo-Branco, I., Cruz-Jesus, F., & Oliveira, T. (2019). “Assessing Industry 4.0 readiness in manufacturing: Evidence for the European Union”. Computers in Industry, 107, 22-32.

Culot, G., Nassimbeni, G., Orzes, G., & Sartor, M. (2020). “Behind the definition of Industry 4.0: Analysis and open questions”. International Journal of Production Economics, 226, 107617.

Du, W. D., Pan, S. L., & Huang, J. (2016).” How a Latecomer Company Used IT to Redeploy Slack Resources”. MIS Quarterly Executive, 15(3).195-213

Frank, A. G., Dalenogare, L. S., & Ayala, N. F. (2019).” Industry 4.0 technologies: Implementation patterns in manufacturing companies”. International Journal of Production Economics, 210, 15-26.

Fritzsche, K., Niehoff, S., & Beier, G. (2018).” Industry 4.0 and climate change—Exploring the science-policy gap”. Sustainability, 10(12), 4511.

Guo, L., & Xu, L. (2021).” The Effects of Digital Transformation on Firm Performance: Evidence from China’s Manufacturing Sector”. Sustainability, (13), 1~18.

Gray, P., El Sawy, O. A., Asper, G., & Thordarson, M. (2015).” Realizing strategic value

through center-edge digital transformation in consumer-centric industries". Revista do CEAM, 3(1), 37.

Hansen, A. M., Kraemmergaard, P., & Mathiassen, L. (2011). "Rapid Adaptation in Digital Transformation: A Participatory Process for Engaging IS and Business Leaders". MIS Quarterly Executive, 10(4).175-185

Hummon, N. P., & Dereian, P. (1989). "Connectivity in a citation network: The development of DNA theory". Social networks, 11(1), 39-63.

Hummon, N. P., Doreian, P., & Freeman, L. C. (1990). "Analyzing the structure of the centrality-productivity literature created between 1948 and 1979". Knowledge, 11(4), 459-480.

Hummon, N. P., & Carley, K. (1993). "Social networks as normal science∗". Social networks, 15(1), 71-106.

Hansen, R., & Sia, S. K. (2015)." Hummel's Digital Transformation Toward Omnichannel Retailing: Key Lessons Learned". MIS Quarterly Executive, 14(2).51-66

Hess, T., Benlian, A., Matt, C., & Wiesböck, F. (2016). "Options for Formulating a Digital Transformation Strategy". MIS Quarterly Executive, 15(2), 123–139.

Hanelt, A., Bohnsack, R., Marz, D., & Marante, C. A. (2021). "A Systematic Review of the Literature on Digital Transformation: Insights and Implications for Strategy and Organizational Change", Journal of Management Studies, 58(5), 1159~1197.

Joshi, A., Benitez, J., Huygh, T., Ruiz, L., & De Haes, S. (2021)." Impact of IT governance process capability on business performance: Theory and empirical evidence". Decision Support Systems, 113668.

Karimi, J., & Walter, Z. (2015). "The role of dynamic capabilities in responding to digital disruption: A factor-based study of the newspaper industry". Journal of Management Information Systems, 32(1), 39-81.

Kauffman, R. J., Li, T., & Van Heck, E. (2010). "Business network-based value creation in electronic commerce". International Journal of Electronic Commerce, 15(1), 113-144.

Kunkel, S., & Matthess, M. (2020)." Digital transformation and environmental sustainability in industry: Putting expectations in Asian and African policies into perspective". Environmental Science & Policy, 112, 318-329.

Lowry, P. B. B., Ali, S., Lai, F. W., Dominic, P. D. D., Tech, V., & Ali, R. F. (2021). "Stock market reactions to favorable and unfavorable information security events: A systematic literature review". Computers & Security, 102451.

Loonam, J., Eaves, S., Kumar, V., & Parry, G. (2018)." Toward digital transformation: Lessons learned from traditional organizations". Strategic Change, 27(2), 101-109.

Liu, J. S., & Lu, L. Y. (2012). "An integrated approach for main path analysis: Development of the Hirsch index as an example". Journal of the American Society for Information Science and Technology, 63(3), 528-542.

Liu, J., Yang, W., & Liu, W. (2021). "Adaptive capacity configurations for the digital transformation: a fuzzy-set analysis of Chinese manufacturing firms". Journal of Organizational Change Management. 34(6), 1122–1241.

Liu, P., & Lu, C. (2021). "Strategic analysis and development plan design on digital transformation in the energy industry: A global perspective". International Journal of Energy Research, 45(14), 19657-19670.

Majchrzak, A., Markus, M. L., & Wareham, J. (2016). "Designing for digital transformation: Lessons for information systems research from the study of ICT and societal challenges". MIS quarterly, 40(2), 267-277.

Maltese, V. (2018)." Digital transformation challenges for universities: Ensuring information consistency across digital services". Cataloging & Classification Quarterly, 56(7), 592-606.

Matarazzo, M., Penco, L., Profumo, G., & Quaglia, R. (2021). "Digital transformation and customer value creation in Made in Italy SMEs: A dynamic capabilities perspective". Journal of Business Research, 123, 642-656.

Menz, M., Kunisch, S., Birkinshaw, J., Collis, D. J., Foss, N. J., Hoskisson, R. E., & Prescott, J. E. (2021). "Corporate Strategy and the Theory of the Firm in the Digital Age". Journal of Management Studies, 58(7), 1695-1720.

Mikheev, A., Serkina, Y., & Vasyaev, A. (2021). "Current trends in the digital transformation of higher education institutions in Russia". Education and Information Technologies, 1-15.

Nicolás-Agustín, Á., Jiménez-Jiménez, D., & Maeso-Fernandez, F. (2021)." The role of human resource practices in the implementation of digital transformation". International Journal of Manpower.

Okhrimenko, I., SOVIK, I., PYANKOVA, S., & LUKYANOVA, A. (2019)." Digital transformation of the socioeconomic system: prospects for digitalization in society". Revista Espacios, 40(38).

Rêgo, B. S., Jayantilal, S., Ferreira, J. J., & Carayannis, E. G. (2021). "Digital Transformation and Strategic Management: a Systematic Review of the Literature". Journal of the Knowledge Economy, 1-28.

Ramírez-Montoya, M. S. (2020)." Transformación digital e innovación educativa en Latinoamérica en el marco del CoVId-19". Campus virtuales, 9(2), 123-139.

Safiullin, M. R., & Akhmetshin, E. M. (2019). "Digital transformation of a university as a factor of ensuring its competitiveness. International Journal of Engineering and Advanced Technology", 9(1), 7387-7390.

Teece, D. J. (2007)." Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance". Strategic management journal, 28(13), 1319-1350.

Tijan, E., Jović, M., Aksentijević, S., & Pucihar, A. (2021). "Digital transformation in the maritime transport sector". Technological Forecasting and Social Change, 170, 120879.

Tsou, H. T., & Chen, J. S. (2021)." How does digital technology usage benefit firm performance?

Digital transformation strategy and organizational innovation as mediators". Technology Analysis & Strategic Management, 1-14.

Vial, G. (2019). "Understanding digital transformation: A review and a research agenda". The journal of strategic information systems, 28(2), 118-144.

Warner, K. S., & Wäger, M. (2019). "Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal". Long Range Planning, 52(3), 326-349.

Yang, Z., Chang, J., Huang, L., & Mardani, A. (2021)." Digital transformation solutions of entrepreneurial SMEs based on an information error-driven T-spherical fuzzy cloud algorithm". International Journal of Information Management, 102384.

Zhu, K., Dong, S., Xu, S. X., & Kraemer, K. L. (2006)." Innovation diffusion in global contexts: determinants of postadoption digital transformation of European companies". European journal of information systems, 15(6), 601-616.

Understanding Covid-19 Impact on Supply Chain through Social Media Discussion

Abstract

The Covid-19 pandemic has created significant disruption to the global supply chain, especially to the food and medical supply chain. This study aims to investigate the impact of Covid-19 on the supply chain using 545,125 supply chain tweets collected between March 2020 and May 2022 from 347,593 unique users globally. The results show that most of the people's sentiment toward the supply chain is neutral in the years 2020 and 2021, and negative sentiment became dominant starting in January 2022. Emotion analysis shows a high level of sadness and optimism, a moderate level of anger at the start of pandemic and the level of anger has increased steadily since June 2021. In addition, topic modeling shows different topics that were discussed in each year. In 2020, the major topics discussed were the government's response to covid-19, the food supply chain crisis, and the medical supply chain crisis. In 2021, the topics moved to inflation/gas price, government handling of the supply chain crisis, and vaccination/recovery. In the first half year of 2022, the dominant topics are inflation and war in Ukraine, inflation and human rights, and US election discussion and bordering crossing issues. In addition, a retweet network was generated to understand how Twitter users interact with each other and how messages travel on social media. Top 10 communities were identified, including the Democrats group, the Republican group, government officials and the news agency group. The result shows that there is great divide/opposite content between the Democrats group and the republican group.

Introduction

The COVID-19 pandemic caused widespread disruptions to global supply chains when many manufacturers in China were shut down due to the virus In January 2020. China is a major supplier of raw materials and intermediate goods to countries around the world, so this disruption had significant consequences. On March 11, 2020, The WHO characterized COVID-19 as a pandemic. California also becomes the first state to order all residents to stay home with the exceptions of going to an essential job or shopping for essential needs. As cases grow, hospitals become overwhelmed, and there is a nationwide shortage of personal protective equipment (PPE). Many food manufacturers were forced to destroy their products due to a lack of buyers, causing disruptions to the food supply chain. The COVID-19 pandemic also caused disruptions to the normal operation of businesses leading to a significant imbalance between supply and demand. This has increased the uncertainty and unpredictability faced by companies and their supply chains (Iftikhar et al., 2021; Tarigan et al., 2021).

Several recent studies have been done to explore the impact of the pandemic on supply chain disruption and beyond. For example, Cariappa et al. (2022) surveyed 729 consumers and 225 farmers in India and found the pandemic-induced lockdowns have hindered access to food markets, and a majority of consumers have experienced price increases leading to food loss along the supply chain and wastage at consumers end. In addition, the lockdown has impeded vegetable farmers' access to markets, thus limiting their productive and sales capacities in Bangladesh (Alan and Khatun, 2021). Exposure to global supply chain disruptions played a

significant role in U.S. cross-industry PPI (Producer Price Index) inflation between January and November 2021 (Statacreu and LaBelle, 2022). The announcement of supply chain disruption events (lockdown, covid-19 case outbreaks, etc.) has also had a negative impact on the stock market price of meat processing companies (Nguemgaing & Sant'Anna ,2021).

Social media has received a lot of attention in recent years from business and research communities. However, the field of supply chain management (SCM) has been slow in studying social media data. There exist only three papers using Twitter data in the supply chain study. Mishra and Singh (2016) used tweets to understand waste in the beef supply chain. They used the keyword "beef" to collect tweets from January 2015 and January 2016, performed sentiment analysis on the tweets, and used the tweets with negative scores to identify major issues posted by consumers about beef. Another study conducted by Chen at al. (2020) collected Twitter data from January 23, 2020 to March 7, 2020 to understand users' discourse and psychological reaction to covid-19 and identified 11 topics, including "updates about confirmed cases," "COVID-19 related death," "cases outside China (worldwide)," "COVID-19 outbreak in South Korea," "early signs of the outbreak in New York," "Diamond Princess cruise," "economic impact," "Preventive measures," "authorities," and "supply chain.". In addition, Chae (2015) proposed a framework (as shown in the below figure) to incorporate tweets in the supply chain study that includes three types of analytics: descriptive analytics (DA), content analytics (CA), and network analytics (NA).

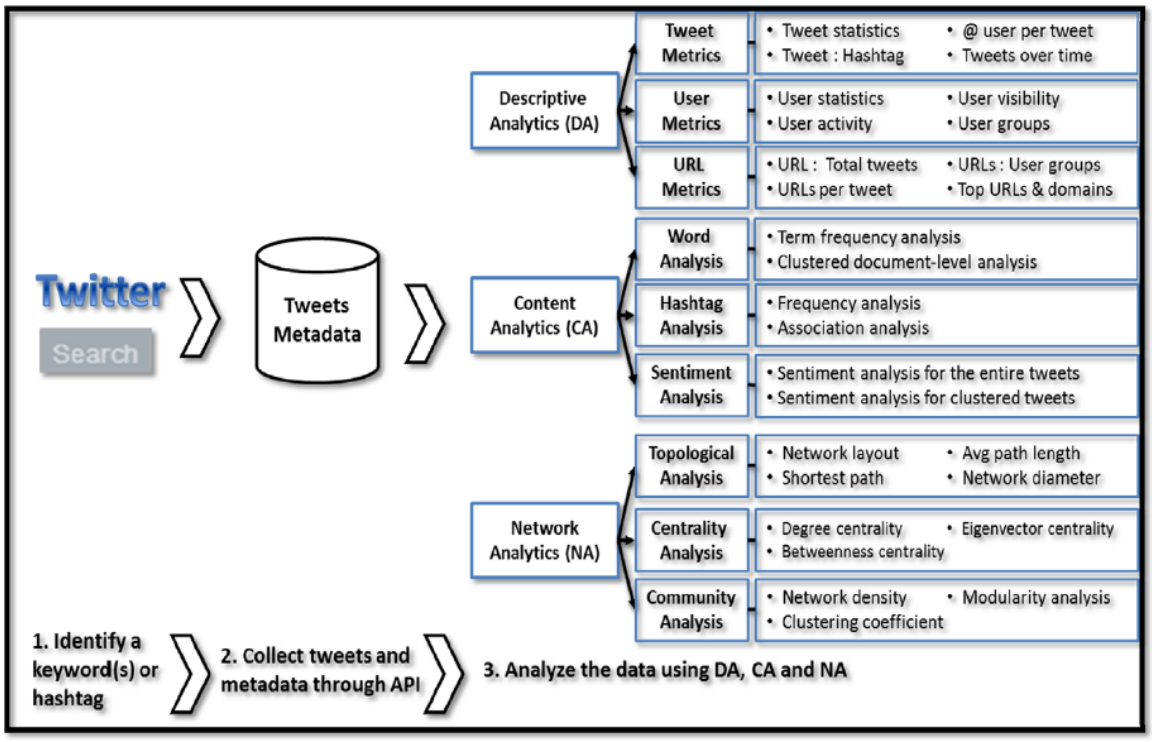


Fig. 1. A proposed framework of extracting intelligence from Twitter (Twitter Analytics).

Following the guideline from the study of Chae (2015) and extending Chen et al. (2020)'s analysis, this study aims to understand the impact of covid-19 on the supply chain using Twitter data by answering the following questions:

- What are the statistics of users and tweets (DA)?
- What are the sentiment and emotions of supply chain tweets (CA)?
- Are there any dominant entities (location, person, organization) mentioned in the tweets(CA)?
- What supply chain topics are shared on Twitter? Are there any prevalent topics? (CA)
- Are there any sub-communities in those Twitter users who are discussing supply chain-related topic during the pandemic? (CA)
- Who are the dominant users in each sub-community (NA)?

Research Methodology

This research is based on covid-19 tweets collected from March 2020 and May 2022 by one of the authors and contained about 1.6 billion tweets. The tweets were filtered using "supply chain" keyword. The total tweets meeting the criteria are 545,125 from 347, 593 unique users. All tweets have included "covid-19 supply chain" in the body of the text, representing social media discussion of the supply chain in the context of the covid-19 pandemic.

The data has a total of 32 columns, displaying relative information about Tweet. This research will mainly focus on the user's name, user location, tweeted text, language used, date of the tweet, and user retweet. Due to the big size of data, the data was stored in an Amzon S3 Storage and analyzed using pyspark, python, and SQL running on a Databricks cluster.

Pre-processing the raw dataset

Raw tweets were pre-processed using pyspark and spark-nlp package. The pre-processing plan was as follows:

1. All words were converted into lowercase.
2. All words starting with http were removed since URLs did not contribute to the text analysis. Hashtags and mentions were since they represent the popularity/trends of particular topics and users.
3. Special characters, punctuations, and numbers were removed from the dataset as they did not help with detecting the profanity comments.
4. Mis-spelling was corrected
5. Lemmatization was applied to switch any word to its base root. For example, stops, stopped, stopping will be converted to stop.
6. All stop words in English were removed
7. Additional words that were used to extract tweets (supply, chain, covid) were removed since all tweets include those words.

Discussion of Data Analysis

This section will analyze the supply chain tweets from three perspectives: descriptive analytics (tweet and user statistics), content analytics (sentiment, emotion, entity detection, and topic modeling), and network analytics (sub-community detection and influencer identification)

Descriptive Analytics

Tweet Statistics

Among 545,125 tweets, retweets account for 71.7% (391,025). There are 16, 299 different hashtags in the tweets, and the top 10 hashtags are '#logistics', '#bbc', '#farmingtoday', '#china', '#pandemic', '#cybersecurity', '#manufacturing', '#inflation', '#blockchain', '#exclusive'. The top hashtags covers a variety of topics, including news media, security, supply chain players, technology, inflation and the pandemic, which have impacted the supply chain. The majority of the tweets (84.3%, 459, 645) do not contain hashtags, about 7.6% (41,232) include one hashtag, 4.1% (22,517) include two hashtags, and the rest 4% include at least three hashtags.

In addition, there are 67,463 mentions , and the top 10 mentions are '@weijia', '@JoJoFromJerz', '@POTUS', '@JoeBiden', '@DonaldJTrumpJr', '@narendramodi', '@CNN', '@Chellaney', '@maddow', and '@ChefGruel']. About 18.8% of tweets have no mentions, 69.9% have one mention, 7.38% have two mentions, and the rest (3.9%) have at least three mentions. It can be seen that most of the top-mentioned accounts are for US government officials. In addition, Narendra Modi (@narendramodi) is the Prime Minister of India, and Brahma Chellaney (@Chellaney) is an author and commentator in India. JoJoFromJerz is a democrat and co-founder of PillowFightCo with 757.8K followers, and @ChefGruel is a chef in California with 174.5K followers.

User Statistics

There are 347, 593 unique users, and each user sends out 1.6 tweets. There are about 2.6% (9,036) of verified users. For the users who provided a location, the top 5 locations of the user are USA, India, England, Canada, and Austria.

About 27.2% have less than 100 followers, 40.6% have between 100 and 1000 followers, 26.5% have between 1000 and 10, 1000 followers. The rest (5.7%) have more than 10,000 followers.

Regarding the accounts a user follows, about 13.8% follow less than 100 users, 48.7% follow between 100 and 1000 users, and 35.2% follow between 1000 and 10, 1000 users. The rest (2.3%) follow more than 10,000 users. In general, most users have more friends than followers.

The top 10 users who tweeted the most are booly_supply, SweetwaterCoFL, OttLegalRebels, sonnycharette, HappyChichester, LTDManagement, CoronaUpdateBot, DBSchenkerUSA, cybersec_feeds, and chidambara09. The top 10 users who received the most retweets are weijia, JoJoFromJerz, DonaldJTrumpJr, narendramodi, CNN, Chellaney, maddow, ChefGruel, ChrisMurphyCT, and EricBoehlert. It is interesting to note that there is no overlap between the top users who tweeted the most and who received the most retweets. In addition, the majority of the users who received the most retweet also have the most mentions.

Content Analytics -Sentiment, Emotion, and Entity Detection

Sentiment, emotion, and entity detection were executed using the package from Hugging Face and John Snow Labs. Hugging Face is a company that provides tools and services for natural language processing (NLP). They are best known for their open-source library of NLP tools, called "Transformers," which provides state-of-the-art machine learning models for a variety of NLP tasks. The Transformers library is built on top of the PyTorch and TensorFlow libraries and is designed to be easy to use and extend. John Snow Labs is a data and technology company that specializes in providing high-quality data products and services to organizations in various industries. Some of the data products and services offered by John Snow Labs include data sets, data integrations, and data consulting. Both companies are the leading providers of the state of art open source NLP packages.

Sentiment analysis was performed using a pre-trained model from Hugging Face (Twitter-roBERTa-base for Sentiment Analysis). This is a roBERTa-base model trained on ~124M tweets from January 2018 to December 2021 and finetuned for sentiment analysis with the TweetEval benchmark. Each tweet is classified with a sentiment label (Positive, Neutral, Negative) with a confidence score.

Similarly, emotion analysis was performed using a pre-trained model from Hugging Face (Twitter-roBERTa-base for Emotion Recognition). This is a roBERTa-base model trained on ~58M tweets and finetuned for emotion recognition with the TweetEval benchmark. Each tweet is classified into four emotions (joy, optimism, anger, and sadness) with a confidence score.

Entity detection is performed using a pre-trained model (bert_token_classifier_ner_btc) from spark-nlp package developed by John Snow Labs. The model identifies three entities (location, person, and organization) mentioned in each tweet.

Sentiment Analysis

Figure 1 shows the monthly number of tweets that were classified as positive, neutral, or negative between March 2020 and May 2022. It can be seen that the majority of tweets are either neutral or negative. There was a spike of negative sentiment on April 2020, then decreased over the years, and it increased dramatically on January 2022.

A word could was created using the top 100 words in April 2020 and January 2022 and is shown in Figure 1a and Figure 1b. It can be seen that the most frequent words in April 2020 are Kushner, @weijia, charge, medical, food, and deliver. Jared Kushner, the former president's son-in-law, was tasked with spearheading the procurement of badly needed medical supplies and personal protective equipment for the nation's medical workers when the pandemic started. He was criticized for his lack of experience and mishandling of the supply chain crisis during that time. In addition, when the pandemic started, people worried about the impact of the pandemic/lockdown on the food supply chain. Below are a few tweets from April 2022:

RT @forwardarc: 'Why is Jared Kushner in charge of anything?' 'What are his qualifications ?' 'Is that why supply chain is screwed up?", "RT @ARiquier: This is a godawful, heartbreaking

situation. Not enough food... but too much food. When "disrupted supply chains" means anima…", "RT @CNN: Tyson Foods is warning that "millions of pounds of meat" will disappear from the supply chain as the coronavirus pandemic pushes", "RT @CNN: With coronavirus severely affecting the potato supply chain, a farm in Idaho is giving away about 2 million potatoes so they don't…", and "RT @ChrisMurphyCT: The federal government is choking supply chains for states and making it even harder to get PPE and medical supplie".

Figure 1b shows that the top words in January 2022 are crisis, omicron, China, border, inflection, and capitalist. As the pandemic is in its third year, people worry about many things, including economy, omicron, inflation and supply chain crisis, and therefore their sentiment on supply chain is worsening. Below are a few tweets from January 2022:

"@POTUS Where? The economy sucks, the supply chain is practically non-existent, COVID is out of control, this country is further divided, and Congress continues to be completely dysfunctional! Where is anything moving forward and/or improving? Ds and Rs = Lying Failures!! https://t.co/r9Uw38Uzv3", "RT @MarkMeadows: The border is a mess. The supply chain is broken. Crime is up in major cities across the country. COVID hospitalizations", " RT @TIME: "Profit doesn't exist anymore." Restaurants that barely survived COVID-19 closures now face labor, inflation and supply chain cri…", "RT @ChefGruel: Is it not possible to make covid test kits in America? We bought 500 million from China during a supply chain crises. Let th…", "Dear @JoeBiden your approval ratings are at 33%. The supply chain is a mess, covid has exploded. Inflation is out of control, your own party has buyers remorse and the supreme court just bitch slapped your mandate policy. Just wanted to give you the attaboy you deserve big guy. https://t.co/ibKSHBOJgU".

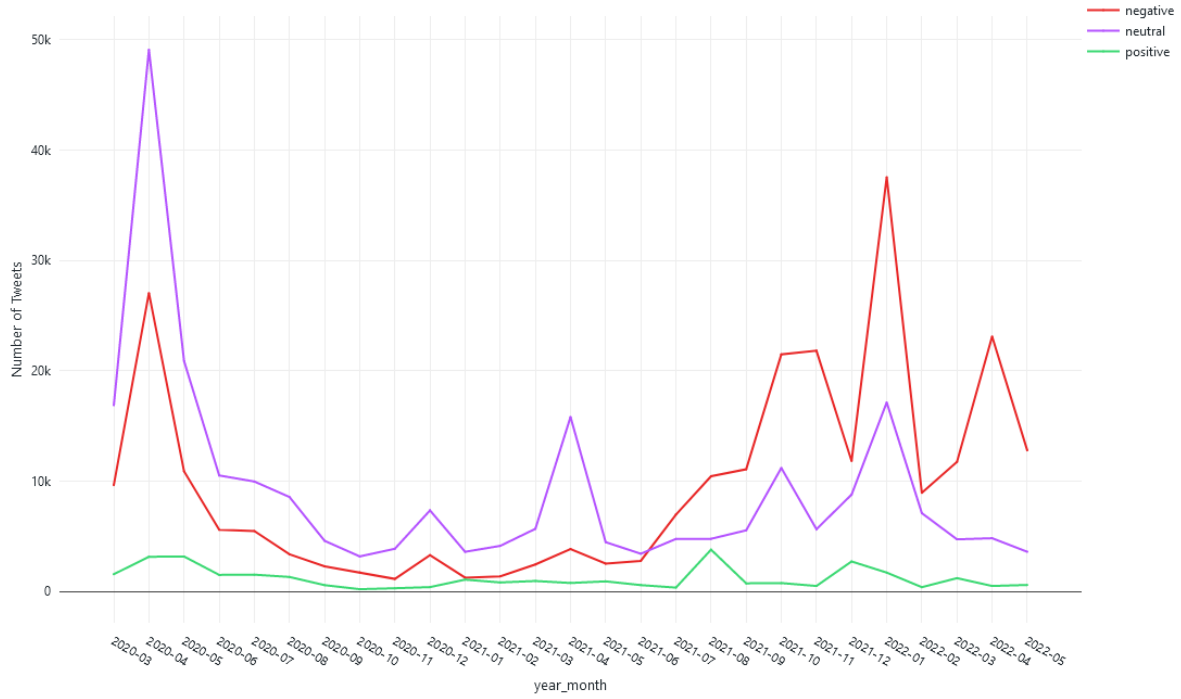


Figure 1. Sentiment of Supply Chain Tweets between March 2020 and May 2022

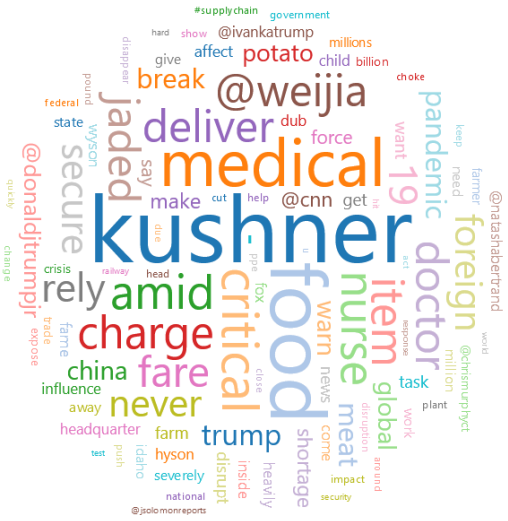


Figure 1a Word cloud in April 2020

Figure 1b Word cloud in January 2022

Emotion Analysis

Figure 2 shows the emotion of supply chain tweets between March 2020 and May 2022. When the pandemic started, the highest emotion is sadness, followed by optimism and anger. Joy has the lowest score. Beginning in June 2021, the level of anger has increased and passed optimism in June 2021 and reached the highest level in April 2022.

Figure 2a shows a word cloud of tweets in April 2022. Crisis is the most mentioned word. Other dominant words include omicron, inflation, capitalist, failure, murder, China, and shortage. The example of the tweets in April 2022 include:

"RT @JoJoFromJerz: Trump didn't win the election. Dr. Fauci didn't create Covid. Biden isn't to blame for the price of gas, the supply chain…", "RT @zerohedge: *SHANGHAI REPORTS FIRST DEATHS DURING CURRENT COVID OUTBREAK China really wants those supply chains to stay broken.", "#GreatResignation, COVID-19, and the war in Ukraine are the most prominent challenges for the food #supplychain this year. @SeanAshcroft from @SupplyChainD reveals the details of the problems and discusses the ways to manage them: https://t.co/WBprYmoFCs", "Supply chain disruptions impacting dairy farmers amid ongoing COVID-related issues" #biznews #businessnews #businesstalk The nation's dairy farmers are feeling the squeeze of the pandemic, as supply chain issues are raising the cost of producing and packaging milk. https://t.co/7A2SLndxkL", "Chinese benchmarks down by 5 percent today(covid related), if supply chains continue to remain disrupted for a few more quarters, inflation could wreak havoc in the short term. Atleast until demand slows down drastically. #China".

It can be seen that people worry about inflation, the Chinese covid-19 policy, the war on Ukraine, and its impact on the supply chain.

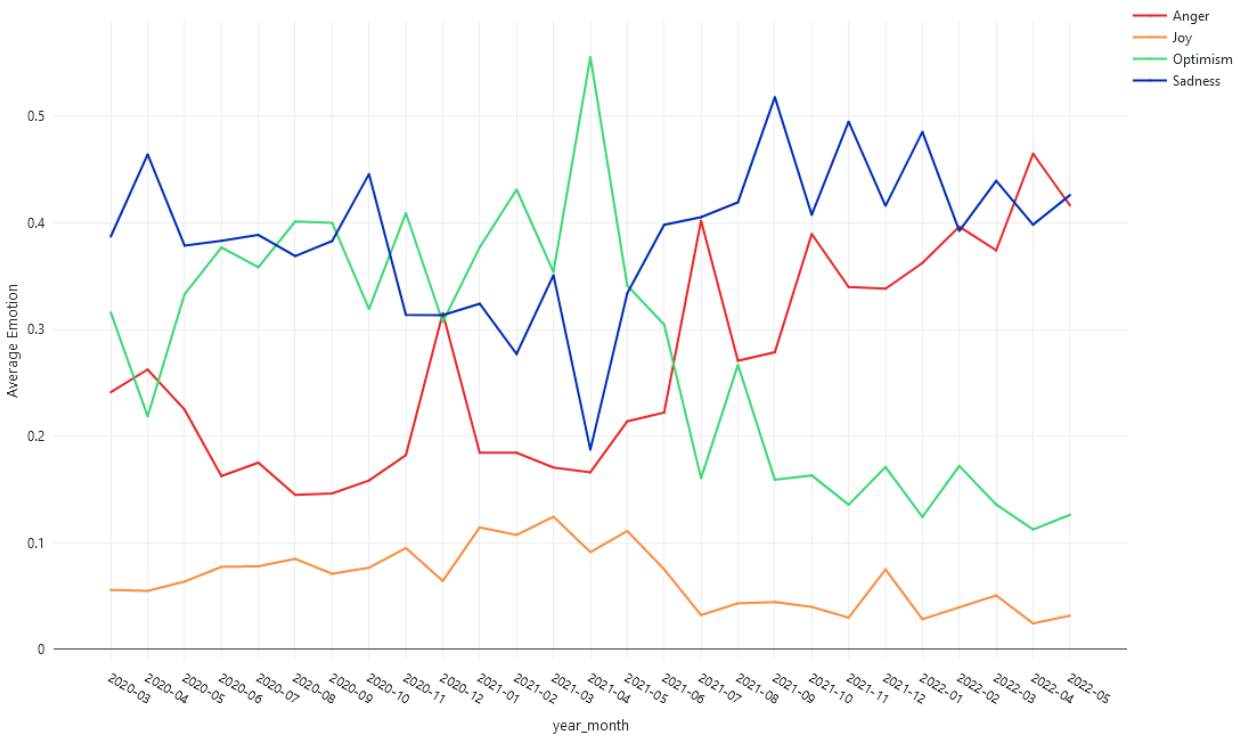


Figure 2 Emotion of Supply Chain Tweets between March 2020 and May 2022

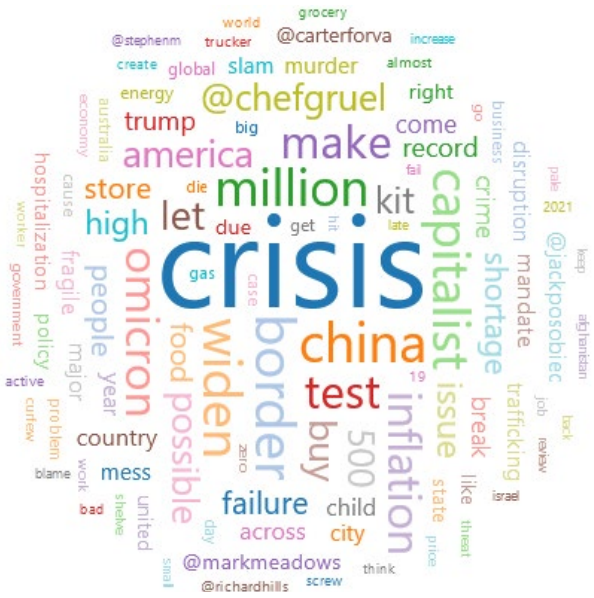


Figure 2a Word Cloud in April 2022

Entities Detections

Figure 3a-3c shows the top locations, persons and organizations detected in tweets. Figure 3a shows that the top 10 mentioned locations are China, America, UK, Afghanistan, European,

Ukraine, Russi, Japan, Indian, and Shanghai, China. The top 10 organizations mentioned are Pfizer, IBM, pentagon, supreme court, Apple, Toyota, Amazon, MSNBC, Tyson Foods, and Mcdonald's.

In addition, the top 5 most mentioned persons are Jared Kusher, Joe Biden, John Polowczyk, Alejandro Mayorkas, and Scott Marrison. John Polowczyk served as the White House Supply Chain lead on the Coronavirus Task Force to ensure healthcare workers get what they need when they need it. President Joe Biden appointed Alejandro Mayorkas as the Secretary of the Department of Homeland Security on February 2, 2021. Scott Morrison is the prime minister of Australia, and he held high-level talks with senior ministers and top department officials on January 2022 to tackle the supply chain and workforce shortage crisis gripping Australia.

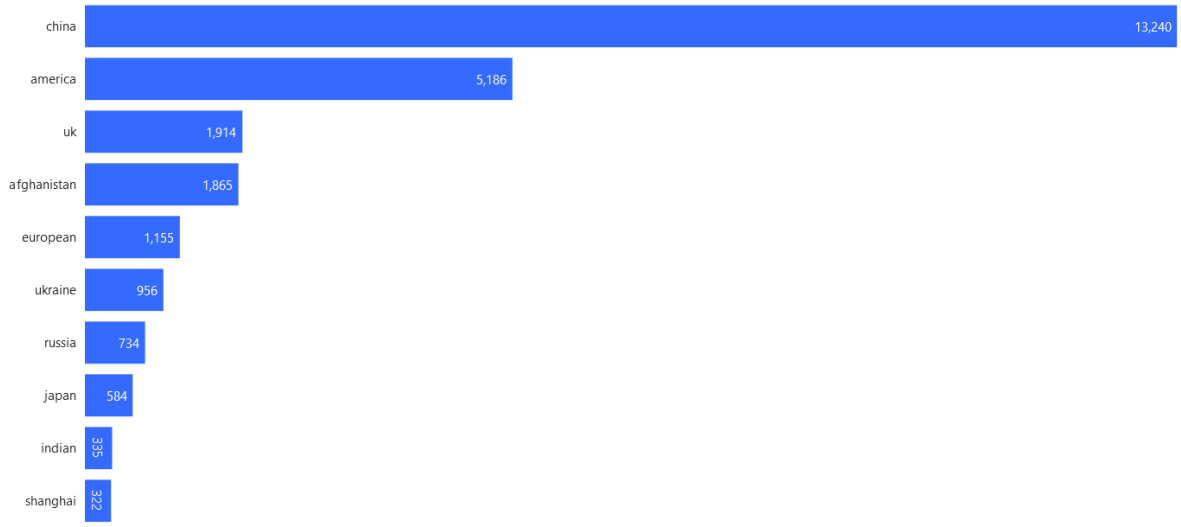


Figure 3a Top 10 Locations

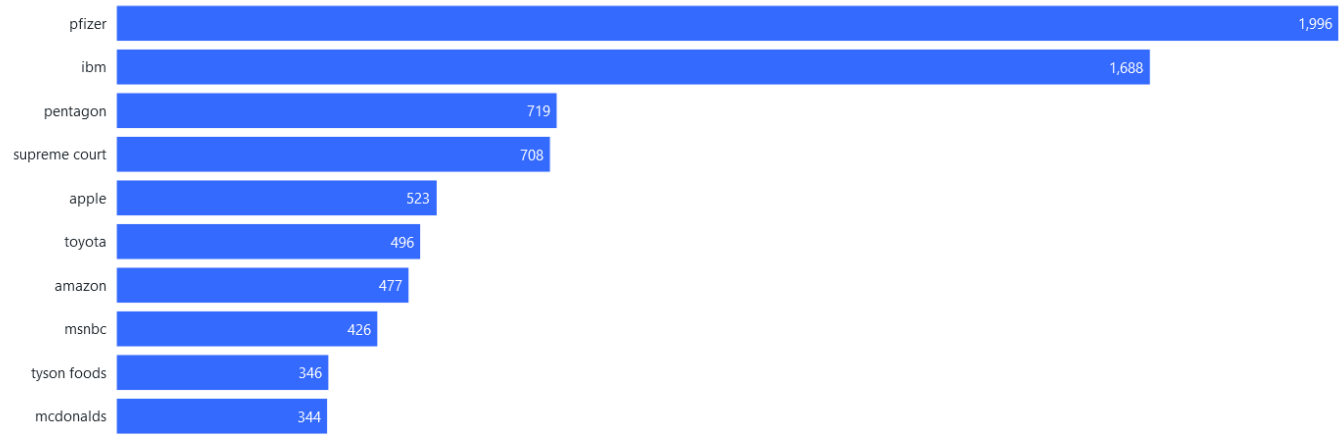


Figure 3b Top 10 organizations

Figure 3c Top 5 Persons

Topic Modelling-Emerging Themes

Topic modeling is an unsupervised machine learning technique that can scan a set of documents to detect words and phrase patterns within them and then cluster word groups and similar expressions that best characterize a set of documents. One popular methodology in topic modeling is Latent Dirichlet Allocation (LDA), a generative probabilistic model that assumes each topic is a matrix over an underlying set of words, and each document is a mixture of a set of topic probabilities.

LDA is used to identify dominant topics in each year, and the results are shown in Table 1 and Figure 4a-4c. Four dominant topics were identified in the year 2020, and three dominant topics were identified in the years 2021 and 2022. Figure 4a shows a visual representation of the four topics in the year 2020. On the left side of the LDA, it shows the four distinct topics. More spreading of the topics indicates more differences between each topic. By clicking on different topics, it can display the distribution of the top 30 words of each possible topic on the right side of the LDA. The blue bars show the overall word frequency in all four topics, while the red bars show the word frequency on the topic selected, which in this case is the top word frequency of topic one in the year 2020.

Table 1 shows the topics identified in each year, the top 30 most relevant terms for each topic, sentiment and emotion of each topic. Four topics emerged in 2020, including government response/policy regarding supply chain, food supply chain crisis, medical supply chain crisis, and potato delivery crisis in Idaho. The dominant emotion in 2020 is sadness. Government response/policy has the highest level of sadness score.

In the year 2021, the three dominant topics are inflation and gas price, US government handling on the supply chain, and vaccination/recovery. Sadness is still the dominant emotion in the first two topics. However, there is a higher level of optimism in topic 3 (vaccination and recovery).

In the year 2022, three dominant topics are inflation and war in Ukraine, inflation and human rights, and US election discussion and bordering crossing issues. The dominant keywords are inflation and crisis. There are a higher level of anger and sadness in those three topics.

Network Analytics-Community Detection

From tweets, we extract the number of retweets received by each user from another user using pyspark. In this network, each node represents one Twitter user, and the edge (the link between two nodes) represents the number of retweets from one user to another.

Label propagation was used to detect communities in networks. Each node in the network is initially assigned to its own community. At every step, nodes send their community affiliation to all neighbors and update their state to the most frequent community affiliation of incoming messages. Label propagation is a standard community detection algorithm for graphs. Table 2 shows the top 10 communities identified through label propagation, it shows the number of users, top 10 retweeted users, top 10 locations, top 10 words, and top 10 hashtags for each community. In addition, Gephi was used to visualize the retweet among users in the top 10 communities, and the result is shown in Figure 5. In Figure 5, nodes represent Twitter users, edges represent the number of tweets one user received from another, and color represents community. Users in the same community have the same color. Figure 5 shows the existence of 10 distinct communities represented by different colors, and Figure 5a-5j shows a visualization of each community using Gephi. It can be seen that most users in the same community are centered around a dominant user and a few are away from the center.

Table 2 shows that two groups (groups 5 and 8) are in India. The dominant users in each group are Democrats (who criticize Trumps), Republicans (who disapprove of Biden), or news agencies who report various issues, including the impact of covid-19, supply chain disruption, gas prices, and inflation.

Conclusions

This study aims to understand the impact of Covid-19 on the supply chain using 545,125 supply chain tweets collected between March 2020 and May 2022 from 347,593 unique users. The results show that people's sentiment toward the supply chain is either neutral or negative in the year 2020, and became more negative starting on January 2022. Emotion analysis shows a high level of sadness and optimism, a moderate level of anger at the start of pandemic and the level of anger has increased steadily since June 2021. In addition, topic modeling shows different topics that were discussed in each year. In 2020, the major topics discussed were the government's response to covid-19, the food supply chain crisis, and the medical supply chain crisis. In 2021, the topics moved to inflation/gas price, government handling of the supply chain crisis, and vaccination/recovery. In the first half year of 2022, the dominant topics are inflation and war in Ukraine, inflation and human rights, and US election discussion and bordering crossing issues. In addition, a retweet network was generated to understand how Twitter users interact with each other and how messages travel on social media. Top 10 communities were identified, including

the Democratic group, the Republican group, government officials and the news agency group. The result shows that there is great divide/opposite content between the Democratic group and the republican group.

References:

Alam GMM, Khatun MN (2021) Impact of COVID-19 on vegetable supply chain and food security: Empirical evidence from Bangladesh. *PLoS ONE* 16(3): e0248120. https://doi.org/10.1371/journal.pone.0248120

Cariappa, AG Adeeth, Acharya , Kamlesh Kumar, Adhav, Chaitanya Ashok Adhav, Sendhil R., Ramasundaram, P. (2022). COVID-19 induced lockdown effects on agricultural commodity prices and consumer behaviour in India – Implications for food loss and waste management. *Socio-Economic Planning Sciences*, 82, 1-23.

Chae, B. (2015). Insights from hashtag #supplychain and Twitter Analytics: Considering Twitter and Twitter data for supply chain practice and research. *International Journal of Production Economic*, 165, 247-259.

Kang, L., Li, Y., Hu, S., Chen, M., Yang, B. (2020). The mental health of medical workers in Wuhan, China dealing with the 2019 novel coronavirus. *The Lancet Psychiatry*, 7, E14.

Mishra, N., Singh (2018), A. "Use of Twitter Data from Waste Minimization in Beef Supply Chain", Big Data Analytics in Operations & Supply Chain Management, 270, 337-359

Santacreu, A. M., LaBelle, J. (2022). Global Supply Chain Disruptions and Inflation During the COVID-19 Pandemic. *Federal Reserve Bank of St. Louis Review*, Second Quarter, 104(2), 78-91.

Iftikhar, A., Purvis, L., & Giannoccaro, I. (2021). A meta-analytical review of antecedents and outcomes of firm resilience. *Journal of Business Research*, 135, 408–425

Nguemgaing, Hélène Flore and Sant'Anna, Ana Claudia (2021), "The Impact of Supply Chain Disruptions on Stock Market Returns during Covid-19". *Faculty & Staff Scholarship*. 3040. https://researchrepository.wvu.edu/faculty_publications/3040

Tarigan, Z. J. H., Siagian, H., & Jie, F. (2021). Impact of Internal Integration, Supply Chain Partnership, Supply Chain Agility, and Supply Chain Resilience on Sustainable Advantage. *Sustainability*, 13(10)

Xue J, Chen J, Chen C, Zheng C, Li S, Zhu T (2020) Public discourse and sentiment during the COVID 19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter. PLoS ONE 15(9): e0239441. https://doi.org/10.1371/journal.pone.0239441

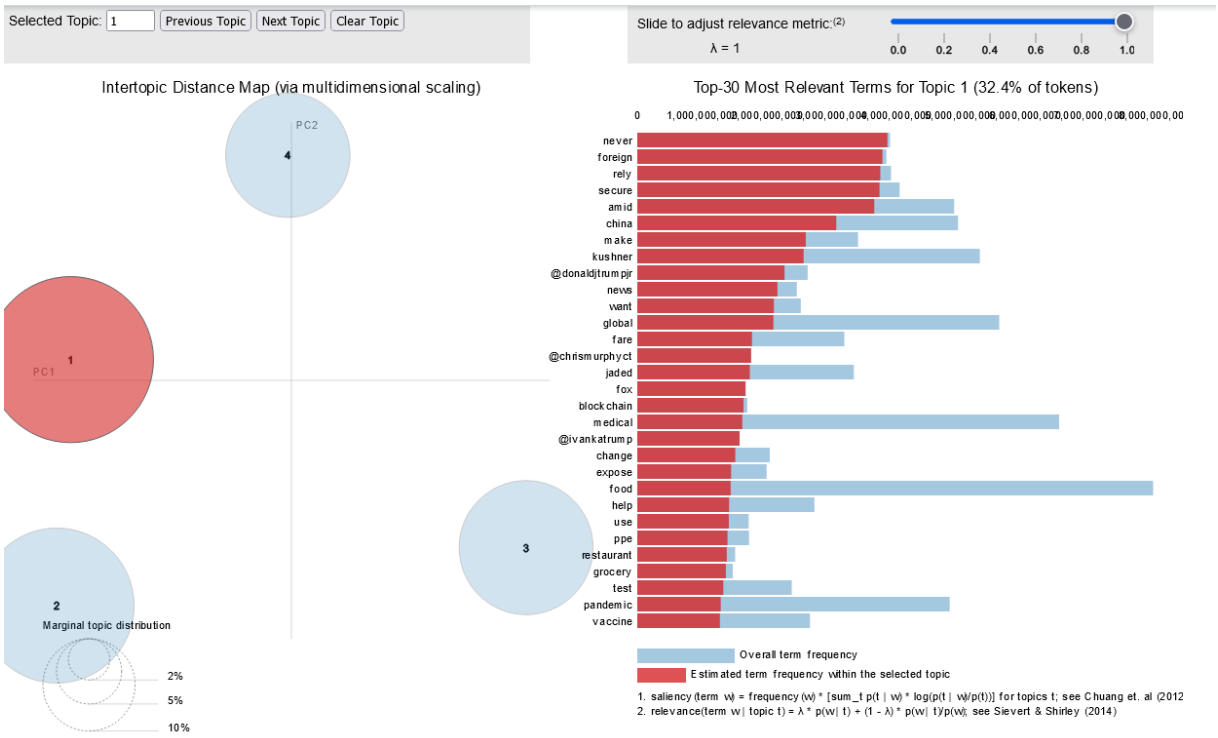


Figure 4a Top Modelling for Supply Chain Tweets in 2020

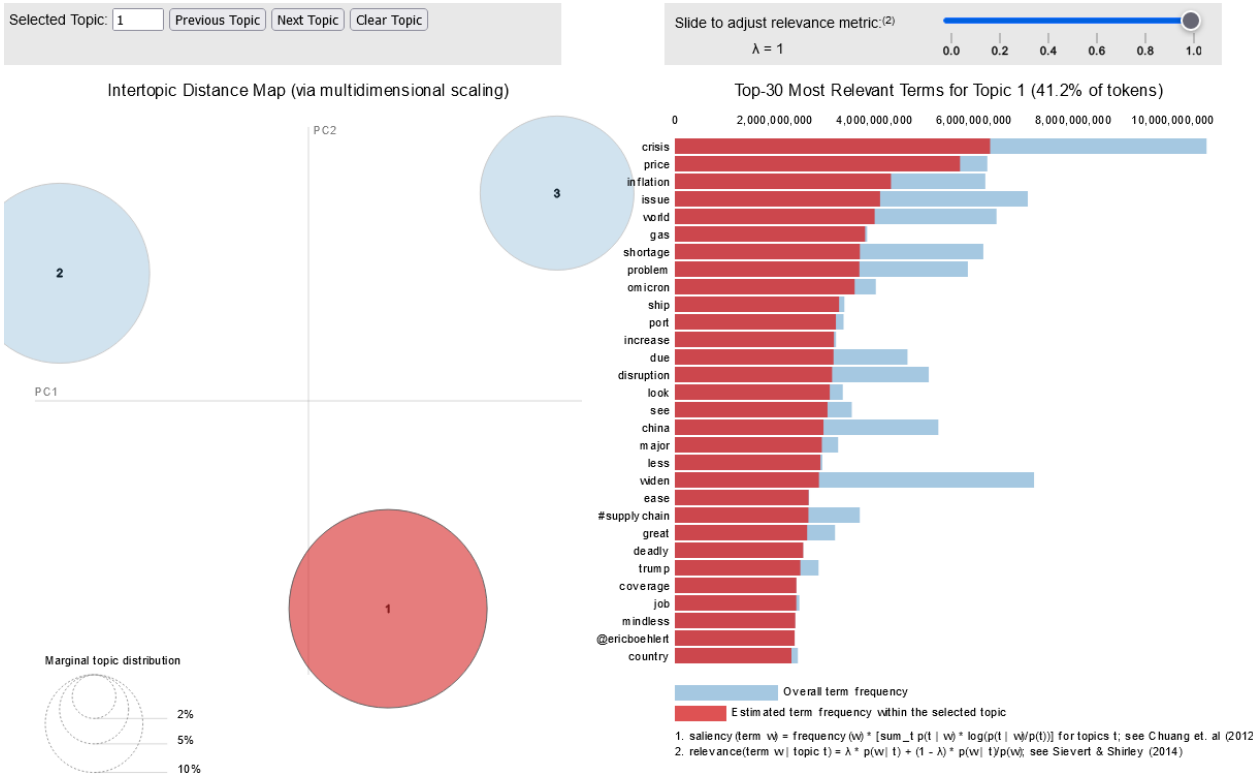


Figure 4b Top Modelling for Supply Chain Tweets in 2021

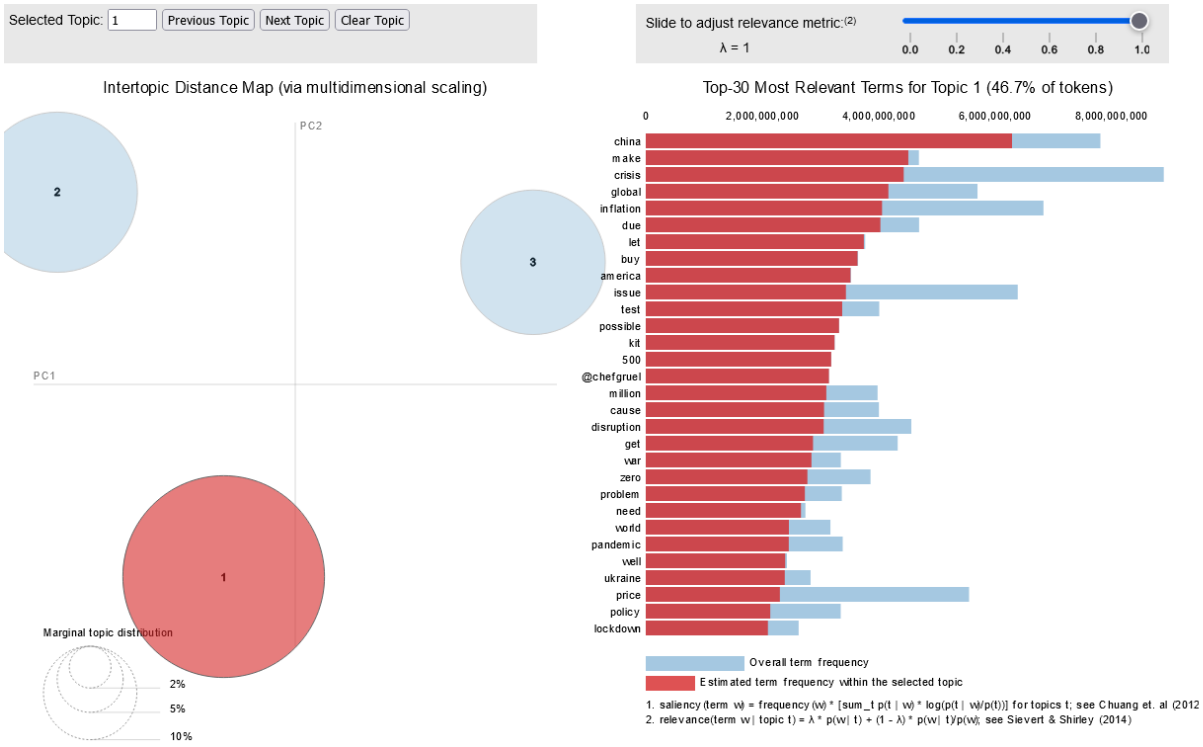


Figure 4c Topic Modelling for Supply Chain Tweets in 2022

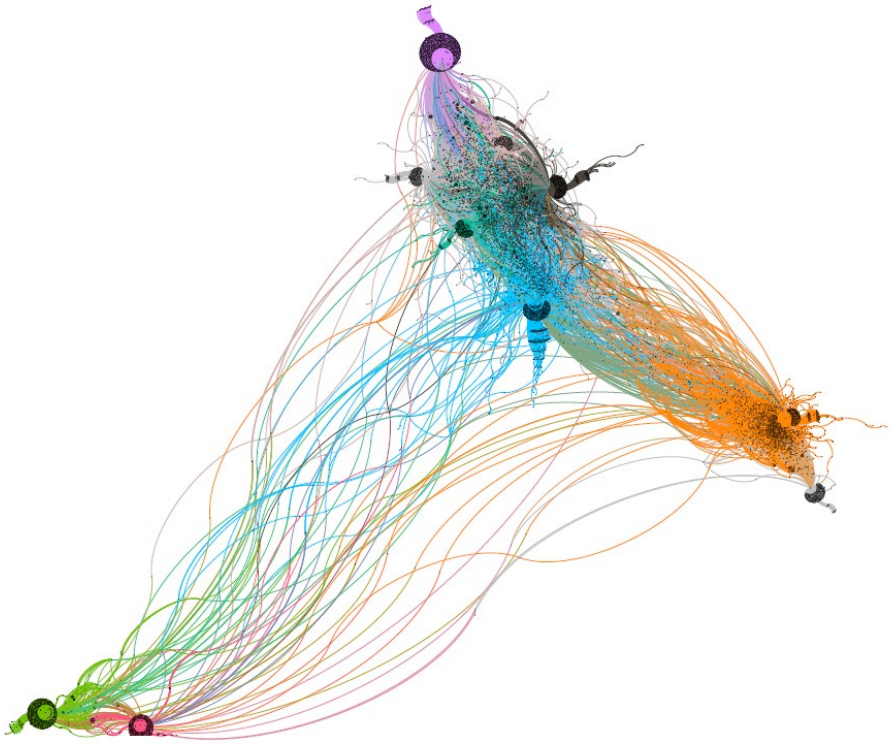


Figure 5 Top 10 Community (created using Gephi)

Table 1. Topic Modelling in Supply Chain between 2020 and 2022

Year	Topics (% of tweets, total tweets)	Top 30 Most Relevant Term for Topic	Sentiment	Emotion
2020	Government response/policy (20.78%, 45,548)	never,foreign,rely,secure,amid,china,make,kushner,@donaldjtrumpjr,news,want,global,fare,@chrismurphyct,jaded,fox,blockchain,medical,@ivankatrump,change,expose,food,help,use,ppe,restaurant,grocery,test,pandemic,vaccine	Positive: 4.77% Neutral: 58.97% Negative: 36.26%	Anger : 0.23 Joy : 0.07 Optimism : 0.27 Sadness : 0.44
	food supply chain crisis (31.9%, 69,976)	food,nationalize,critical,national,meat,medical,pandemic,warn,local,break,say,shortage,global,#supplychain,@maddow,farmer,disruption,disrupt,wyson,trade,#covid19,due,hyson,risk,need,come,security,impact,business,billion	Positive: 8.73% Neutral: 58.97% Negative: 32.31%	Anger : 0.18 Joy : 0.06 Optimism : 0.35 Sadness : 0.41
	Medical supply chain crisis (18,2%, 39,936)	kushner,jaded,deliver,charge,doctor,item,@weijia,nurse,task,force,critical,expert,medical,fame,trump,inside,influence,heavily,dub,child,headquarter,@natashabertrand,vaccine,president,scientist,food,consult,state,long,stock	Positive: 5.84% Neutral: 66.03% Negative: 28.12%	Anger : 0.26 Joy : 0.09 Optimism : 0.28 Sadness : 0.37
	potato delivery crisis in Idaho (29.1%, 63,748)	potato,farm,fare,china,nurse,medical,item,deliver,charge,doctor,@weijia,affect,critical,end,Kushner,million,give,white,house,americana,buy,reliance,global,confirm,severely,executive,away,idaho,order,work	Positive: 4.98% Neutral: 63.34% Negative: 31.67%	Anger : 0.25 Joy : 0.05 Optimism : 0.29 Sadness : 0.41
2021	inflation and gas price (27.4%, 51,966)	crisis,price,inflation,issue,world,gas,shortage,problem,omicron,ship,por,tincrease,due,disruption,look,see,china,major,less,widen,ease,#supplychain,great,deadly,trump,coverage,job,mindless,@ericboehlert,country	Positive: 4.06% Neutral: 26.2% Negative: 69.74%	Anger : 0.32 Joy : 0.04 Optimism : 0.16 Sadness : 0.48

	US government handle on supply chain crisis (42.8%, 81,346)	vaccine,also,biden,crisis,raw,potus,importance,smooth,global,discussion,efficient,@narendramodi,underscore,thank,@joebiden,build,hard,many,nation,lesson,agree,quad,crucial,today,reliant,@chellaney,sacrifice,learn,pandemic,china	Positive: 4.55% Neutral: 41.93% Negative: 53.52%	Anger : 0.31 Joy : 0.06 Optimism : 0.21 Sadness : 0.42
	Vaccination and recovery (29.8%, 56,559)	issue,say,disrupt,problem,even,shortage,cause,get,american,global,think,indie,big,food,world,recovery,vaxxed,economic,failure,midst,historic,around,really,president,yesterday,nation's,@scottadamssays,fac,labour,crisis	Positive: 14.63% Neutral: 53.3% Negative: 32.07%	Anger : 0.24 Joy : 0.08 Optimism : 0.39 Sadness : 0.29
2022	Inflation and war on Ukraine	china,make,crisis,global,inflation,due,let,buy,america,issue,test,possible,kit,500,@chefgruel,million,cause,disruption,get,war,zero,problem,need,world,pandemic,well,ukraine,price,policy,lockdown	Positive: 2.93% Neutral: 30.2% Negative: 66.87%	Anger : 0.38 Joy : 0.03 Optimism : 0.15 Sadness : 0.44
	Inflation and human rights	crisis,capitalist,say,people,issue,border,omicron,high,come,work,inflation,many,store,record,failure,report,biden,walmart,neighbor,son,@ginnymcdonald8,first,get,murder,year,slam,@jackposobiec,trafficking,united,child	Positive: 3.36% Neutral: 35.39% Negative: 61.25%	Anger : 0.35 Joy : 0.03 Optimism : 0.11 Sadness : 0.5
	US election and bordering crossing issues	didn,create,price,blame,election,win,isn,trump,gas,farci,@jojofromjerz,biden,country,city,across,mess,shortage,major,border,trucker,crime,global,break,hospitalization,@markmeadows,open,canadian,policy,food,china	Positive: 3.85% Neutral: 14.65% Negative: 81.5%	Anger : 0.47 Joy : 0.03 Optimism : 0.11 Sadness : 0.39

Table 2. Top 10 Community among Twitter Users

Community	Total Users	Top 10 retweeted users	Top 10 locations	Top 10 words	Top 10 hashtags	Summary
1	9622	["JoJoFromJerz","weijia","RepAdamSchiff","NatashaBertrand","fredifredae","RepAOC","BaddCompani","DelPercioS","VABVOX","SecretaryPete"]	["United States","USA","California, USA","Florida, USA","Texas, USA","New Jersey, USA","Chicago, IL","New York, USA","Los Angeles, CA","Canada"]	["didnt","trump","price","blame","gas","isnt","create","win","electiondr","covidbiden"]	["#inflation","#karenconvoy1","#exclusive","#donthecon","#sad","#capitalism","#n95mask","#gopbetrayedamerica","#ppe","#blockchain"]	Democrats criticize Trumps and his claim about election results
2	8768	["CNN","weijia","Reuters","ChrisMurphyCT","maddow","MSNBC","business","NatashaBertrand","TIME","JoeBiden"]	["United States","USA","California, USA","Texas, USA","Canada","New York, NY","Florida, USA","New York, USA","Earth","Washington, DC"]	["potato","food","affect","farm","million","give","dont","away","severely","idaho"]	["#inn","#china","#fishing","#breaking","#vaccine","#resilien","#inflation","#foodsafety","#iowa","#nebraska"]	News agency talks about potato crisis in Idaho and food supply chain
3	6487	["ScottAdamsSays","EmeraldRobinson","StephenM","ChefGruel","JackPosobiec","MarshaBlackburn","mtgreenee","GOPChairwoman","MarkMeadows","JakeEvansGA"]	["United States","USA","India ðŸ‡®ðŸ‡³","California, USA","Florida, USA","Texas","Texas, USA","Virginia, USA","Arizona, USA","39.109 -76.77"]	["biden","issue","cause","problem","think","president","big","nation","learn","fac"]	["#inflation","#gas","#supplychaincrisis","#toryincompetence","#ai","#data","#chinas","#artificialintelligence","#china","#machinelearning"]	A group who criticizes Biden's handling of Supply Chain Crisis.

4	6260	["maddow","weijia","ChrisMurphyCT","NatashaBertrand","JoJoFromJerz","EricBoehlert","MSNBC","JoeBiden","GinnyMcDonald8","WhiteHouse"]	["United States","California, USA","USA","Florida, USA","Los Angeles, CA","Texas, USA","Pennsylvania, USA","New York, NY","California","Ohio"]	["medical","critical","nationalize","suppliesnationalize","suppliesnationali","kushner","jared","trump","price","charge"]	["#inn","#exclusive","#morrisonmu","#farmingtoday","#protectnurses","#bbc","#pe","#fieldhospitals","#inflation","#stayathome"]	News agency discussing medical supply chain crisis
5	5466	["narendramodi","Chellaney","ANI","DrSJaishankar","nsitharaman","PiyushGoyalOffc","SandhuTaranjitS","BJP4Gujarat","ani_digital","knowthenation"]	["India","New Delhi, India","Mumbai, India","à¤-à¤¾à¤°à¤¤","Ahmadabad City, India","Hyderabad, India","Gujarat, India","Kolkata, India","Jaipur, India","Bengaluru, India"]	["vaccine","underscore","smooth","importance","raw","also","discussion","efficient","build","nation"]	["#indopacific","#straighttalkwithnirmala","#tripswaiver","#foreignsecretaryharshshringla","#drug","#china1","#supply","#g20","#economy","#aftercovid"]	A group in India including prime minister talking about covid policy and drug supply chain
6	4006	["ChrisMurphyCT","weijia","NatashaBertrand","JoJoFromJerz","EricBoehlert","GinnyMcDonald8","MSNBC","VABVOX","RepAOC","MitchLandrieu"]	["United States","California, USA","USA","New York, NY","Connecticut, USA","Florida, USA","Texas, USA","Chicago, IL","New Haven, CT","Los Angeles, CA"]	["medical","critical","test","break","mask","ventilator","include","act","make","emergency"]	["#protectnurses","#whereistheppe","#inflation","#canada","#healthcare","#blockade","#pandemic","#pricegouging","#newsbreak","#60minu"]	Government official discussing medical supply chain

7	3718	["JoeBiden","EricBoehlert","JoJoFromJerz","weijia","WhiteHouse","GinnyMcDonald8","ChrisMurphyCT","RepAdamSchiff","MSNBC","NatashaBertrand"]	["United States","California, USA","USA","Texas, USA","Florida, USA","Los Angeles, CA","New Jersey, USA","Washington, DC","Virginia, USA","Ohio, USA"]	["economic","recovery","world","american","disrupt","around","historic","midst","biden","price"]	["#weo","#chinas","#biden","#supplychainc","#trump","#inflation","#walledout","#walledin","#cancel trumpism","#emptyshelvestrumpstress"]	Joe Biden and white house discuss economic recovery
8	3463	["Chellaney","ANI","PranavDixit","wef","DrSJaishankar","NavroopSingh_","AlexeiArora","startupindia","Kanchan Gupta","ChefGruel"]	["India","New Delhi, India","Mumbai, India","Mumbai","Bengaluru, India","New Delhi","Bangalore","à¤-à¤¾à¤°à¤¤","Pune, India","Bharat"]	["build","learn","many","lesson","nation","thank","hard","agree","quad","chinareliant"]	["#indopacific","#india","#drugs","#operations","#taiwan","#china","#generic","#us","#breaking","#oneyearoflockdown"]	A group in India talking about China's lockdown and Indo-Pacific region
9	3265	["EricBoehlert","JoJoFromJerz","weijia","WhiteHouse","NatashaBertrand","keithedwards","fredifredae","RepAdamSchiff","MSNBC","EricMGarcia"]	["United States","Los Angeles, CA","USA","California, USA","Florida, USA","Seattle, WA","Chicago, IL","Washington, USA","Portland, OR","New York, NY"]	["price","gas","biden","crisis","omicron","problem","deadly","less","ease","look"]	["#protectnurses","#uv","#pandemic","#msnbc","#cnn","#enough","#sm","#bif","#supplychainrecovery","#whereistheppe"]	Government offices talk about Omicron, gas price and personal protective equipment (ppe)
10	2554	["ChefGruel","GOPChairwoman","EmeraldRobinson","DonaldJTrumpJr","erichhartmann","Lancegooden","elonmusk","pvtjokerus","Phaedru96690361","Ivanka Trump"]	["United States","USA","Florida, USA","California, USA","Texas, USA","Ohio, USA","North Carolina, USA","Florida","Alabama, USA","Washington, USA"]	["crisis","china","test","buy","let","make","america","500","kit","possible"]	["#emptyshelvesbiden","#letsgobrandon"]	A group of Republicans talks about Biden's economic and China' policy.

Table 3. Visualization of Each Community using Gephi

Figure 5a (center user: JojoFromJerz)

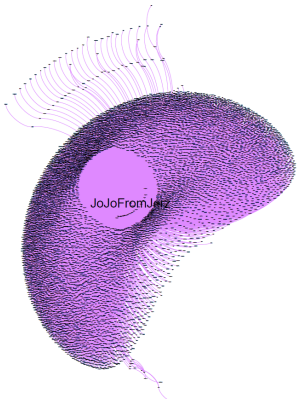


Figure 5b (center user: Narendra Modi)

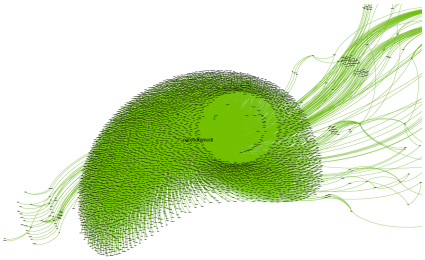


Figure 5c (center user: CNN)

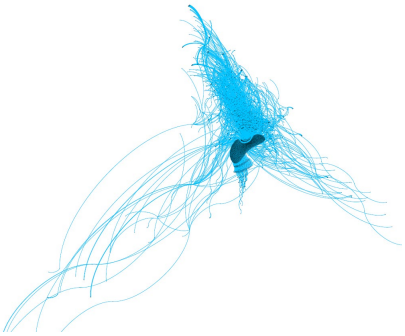


Figure 5d (center user: Rachel Maddow)

Figure 5e (center user: Scott Adam)

Figure 5f (center user: Brahma Chellaney in India)

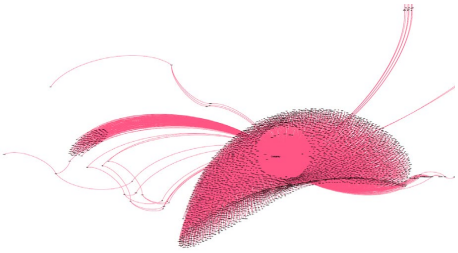


Figure 5g (center user: Chris Murphy, Senator from Connecticut)

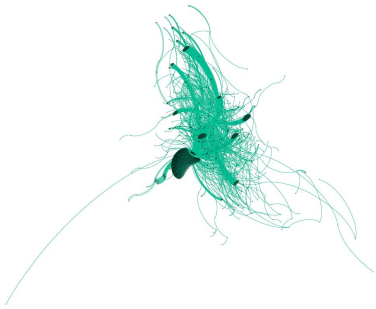


Figure 5h (center user: Joe Biden)

Figure 5i (center user: Eric Boehlert, a journalist, writer, and media critic)

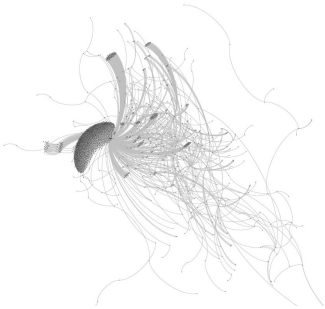
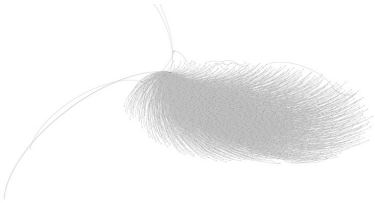


Figure 5j (center user: Chef Andrew Gruel from California)



Cyber Security, IT, and Emerging Technologies

Organizational (Digital) Response to Turbulent Business Environment due to the Pandemic

Abstract

Since 2020, the COVID-19 pandemic has caused major disruptions in all aspects of life due to individuals, communities, and organizations experiencing radical shifts in culture which spurred the adoption of agile approaches with regards to decision-making (Mooney et al., 2020). Digital technologies played a critical role and various digital platforms and technologies have enabled organizations to adapt during the pandemic. Gudi and Chinta (2020) presented empirical evidence that showed awareness of digital capabilities impacts digital entrepreneurship. In this paper, we present a conceptual model that extends their study to include awareness of the pandemic as a force for change.

With the rapid development of digital technologies, we see the emergence of digital entrepreneurship which is the intersection of digital technology and entrepreneurship. This study seeks to investigate the *enabling* role of environmental change for entrepreneurial pursuits, particularly *digital startups*. We aim to develop empirically validated theories focused on environmental change by COVID-19 and digital entrepreneurship. In our research framework, we propose factors related to digital capabilities, the awareness of the consequences due to a pandemic, and how one would pursue implementation options as they strive to succeed in the entrepreneurial world. This quantitative study will use Qualtrics to develop a survey for data collection. Data analyses will be conducted using data analytical methodologies to test the hypotheses, including the moderating and mediating effects of the variables in our research. Our findings will provide impetus for future research studies that explore digital entrepreneurship, especially during crisis situations.

Background and Significance

Since the beginning of 2020, the COVID-19 pandemic has impacted all aspects of life and caused major disruptions in commerce, healthcare, education, government agencies, private enterprises, and many other institutions. Individuals, communities, and organizations have experienced radical shifts in culture and are seeking agile approaches in decision-making tactics (Mooney et al., 2020). Such transformation will require bringing together the expertise from several different sources such as public health officials, immunologists, data scientists, ethicists, business leaders and government policy makers.

Digital technologies have played a critical role during the pandemic. Health agency workers in general and particularly frontline health professionals rely on innovative digital technology solutions so they can provide reliable and effective service to patients (Rai, 2020). As per the Centers for Disease Control and Prevention (CDC), data modernization is a priority to ensure real-time data is available for emergencies (CDC, 2020). For example, the corona disease reports from the electronic health record systems of all 6,100 U.S. hospitals should be automatically made available real-time to public health. Novel mobile technologies are offering contact tracing apps (Contact Tracing Apps, 2020) that lead to choices in adopting technology design related to 5G mobile infrastructures, GPS, and Bluetooth. Digital platforms and technologies have enabled organizations to adapt (e.g. transition to remote work, implement

virtual contactless services) and educational institutions to transition to online and hybrid course delivery options.

With the rapid development of digital technologies, we see the emergence of digital entrepreneurship which is the intersection of digital technology and entrepreneurship (Recker & von Briel, 2019; Gudi & Chinta, 2020). Environmental changes (such as a pandemic) are important sources of stimuli to entrepreneurs in addition to human agency (Shane, 2003), social networks (Alvarez & Barney, 2007; Wood & McKinley, 2010), and current trends in the external world. Though there is an abundance of concepts and statements pertaining to entrepreneurial agents, the strategic role of pandemics remains under-emphasized and under-theorized in contemporary entrepreneurship and strategy research (Agarwal et al., 2017; Chandra, 2018; Davidsson, 2019). The use of digital technologies in many recent entrepreneurial successes highlights the need for increased attention to the role of agent-independent, external change for strategy (Porter & Heppelman, 2014), and entrepreneurship (Autio et al., 2018; Nambisan, 2017). Likewise, the impact of COVID-19 provides a new stimulus for digital entrepreneurship.

A substantial body of theories, concepts, and evidence exists for the entrepreneurial inclinations and capacities of entrepreneurial agents such as individuals and organizations. An equally rich set of theories, concepts, and evidence need to be developed for capturing the entrepreneurial potential stimulated by environmental change caused by pandemics. For these reasons, we aim to develop empirically validated theories focused on environmental change by COVID-19 and digital entrepreneurship.

Our conceptual model seeks to investigate the *enabling* role of environmental change for entrepreneurial pursuits, particularly *digital startups*. Researchers have proposed the particular, *venture-level benefits* strategically and fortuitously derived from environmental changes caused by pandemics. Such benefits may be conceptualized as 'affordances' (Autio et al., 2018; Gibson, 1977); 'enabling mechanisms' (Davidsson et al., 2018); 'unactualized propensities' (Ramoglou & Tsang, 2017), 'autonomous agency' (Brynjolfson & McAffee, 2014) or otherwise.

We consider the following research questions:

- How will entrepreneurs transform organizations and the nature of work through the experiences with digital technologies during the pandemic in conducting remote work, scaling virtual services, and deploying disruptive innovations?
- How can information and communication technologies be developed and deployed to enhance decision-making processes during a pandemic?
- What are the salient commonalities and differences across types and instances of environmental change caused by pandemics regarding the potential they offer for digital entrepreneurship? What precise benefits can be derived from what kinds of change due to digital entrepreneurship and by what types of ventures when pandemics overtake economies?
- Which of these benefits are under- and over-estimated by novice vs. expert entrepreneurs, and by entrepreneurs vs. their investors, respectively?

Although a substantial body of research is available on the risks and uncertainties of start-up businesses, there is a lack of literature on how these would apply to digital entrepreneurship, particularly in a pandemic situation. This relates to understanding the motivation and the

determinants for an entrepreneur to start a new business. Exploring some of the determinants of entrepreneurial drive which are required for digital start-ups is essential. Equally important, is understanding the factors related to digital capabilities, the awareness of the consequences due to a pandemic, and how one would pursue implementation options as they strive to succeed in the entrepreneurial world.

Although it is broadly acknowledged that digital technologies can have significant effects on entrepreneurial processes, most of the existing research in technology entrepreneurship is mainly focused on doing empirical work in technology-intensive environments (Beckman, Eisenhardt, Kotha, Meyer, & Rajagopalan, 2012; Zupic, 2014). For example, SMACIT refers to technologies using social, mobile, analytics, and the Internet of Things (Vial, 2019). In addition to viewing technology as merely a context for entrepreneurship (Bingham & Haleblian, 2012; Vissa & Bhagavatula, 2012), the research spectrum must be extended to include concepts related to digital technologies which influence entrepreneurial drive. One of the motivations for this proposal stems from the fact that research scholars (e.g. Nambisan, 2017) are pointing out that more work needs to be done on specific aspects of digital technologies which can shape "entrepreneurial opportunities, decisions, actions, and outcomes." (p. 1030)

Recker and von Briel (2019) propose a framework wherein digital aspects are assigned to three dimensions: as **enabler** (digital enablers of entrepreneurial processes), **outcome** (digital outcomes of entrepreneurial processes), or **context** (in which entrepreneurial processes take place). Digital technologies that act as enablers have the potential to *change* and *influence* entrepreneurial action, processes, and outcome (Davidsson et al., 2018; von Briel, Davidsson, et al., 2018; Huang et al., 2017; Nambisan, 2017). Digital technologies act as an outcome of entrepreneurial pursuits when they are *created* as the service or product of their market offering (von Briel, Recker, et al., 2018; Lyytinen et al., 2016). The third dimension is concerned with digital technologies that can influence broader *contexts* of entrepreneurship or establish digital ecosystems wherein entrepreneurial processes are created ((Autio et al., 2018; von Briel, Davidsson, et al., 2018). One of the significant outcomes of our proposed research will be to categorize the entrepreneurial processes in one or more of the above dimensions.

Research on resilience theory has shed light on the interaction between crisis and resilience and how organizations can prepare, adapt, and respond to a crisis situation (Williams et al., 2017). Understanding factors related to entrepreneurial drive and motivation in the digital age can open new arenas in theory-based research in resilience and adaptation. Since digitization is not about one technology, our proposal is not intended to be an exhaustive study of digital capabilities; so it is not a roadmap to show how to start a digital venture. Instead our approach is emblematic and is meant to direct attention to how these factors situated in a pandemic frame of reference can influence digital entrepreneurial drive.

Existing literature revealed that environmental changes lead to increased *rates* of new venture creation. Several studies examine effects of 'characteristics of opportunities' where those characteristics tend to be concrete and context-specific rather than expressed as generalizable abstractions. Our proposal attempts to develop strong conceptualization pertaining to external changes and their effects on digital ventures. Amit and Zott (2001), Barreto (2012), Grégoire and Shepherd (2012), Hiatt et al. (2009), and von Briel et al. (2018) are examples of studies that demonstrate such qualities and hence undergird our research proposal. These studies also

illustrate some of the many different theoretical perspectives and approaches that can be fruitfully applied to research on effects of pandemics. The purpose of our proposal is to stimulate, connect, and extend research of this kind to help make research on "Environmental Change by COVID-19 and Digital Entrepreneurship" an integral and growing part of the core of digital entrepreneurship scholarship.

In the realm of digital entrepreneurship, digital entrepreneurial responses are preceded by an awareness of external events, the level of resident digital capabilities, an understanding of context and enablers and inherent motivation to respond. We further conceptualize that the motivation to respond digitally leads to digital entrepreneurial responses and that motivation itself is an intermediate result of awareness of an external stimulus and resident digital capabilities. Thus, our research framework that we will test empirically is depicted in Figure 1 below.

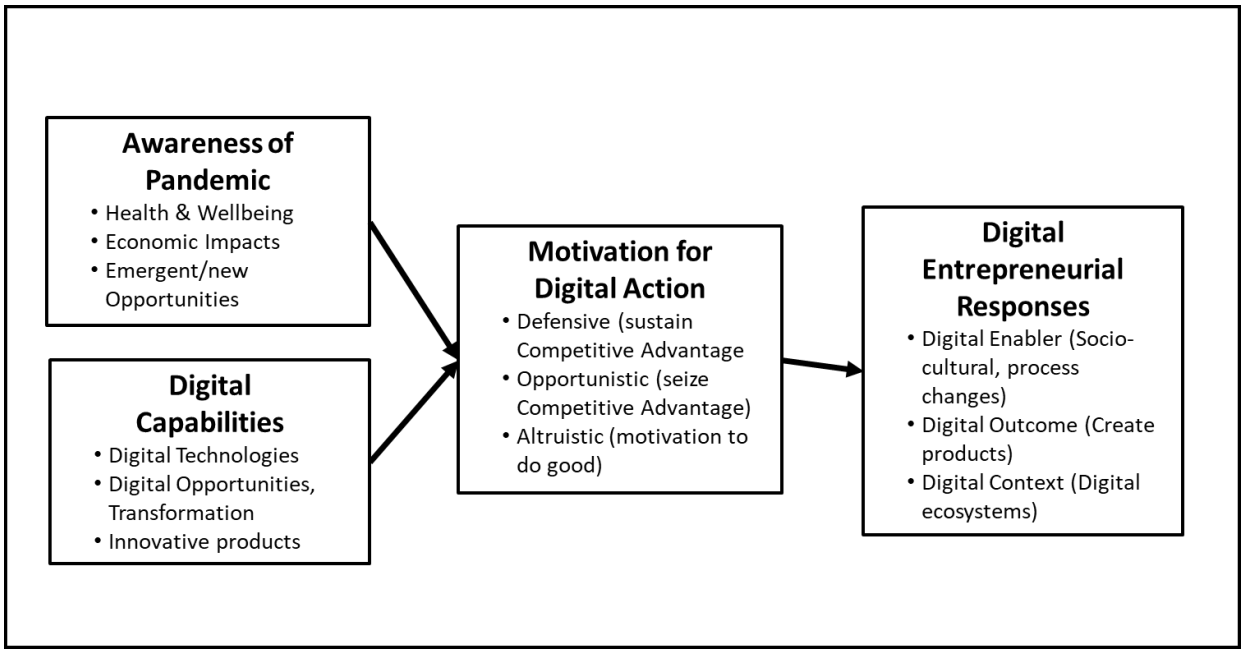


Figure 1: Research Framework

Stemming from the research framework in Figure 1, the following hypotheses emerge:

H_1: Motivation for digital action would be positively impacted by the degree of awareness of the pandemic.

H_2: Motivation for digital action would be positively impacted by the degree of the digital capabilities of the organization.

H_3: Motivation for digital action would positively impact the digital responses by the organization.

H_4 and H_5: Motivation for digital action would mediate the direct effects of awareness of pandemic and digital capabilities on the digital responses by the organization.

Current stage of the research

A competitive research grant has been awarded for this research project in summer 2022. Currently a survey instrument is being constructed from various sources in the extant literature. The questionnaire will be pre-tested to ascertain the validity of the instrument. It is proposed that the data collection will begin in Spring 2023 and the empirical paper will be written in summer 2023 for submission to a peer reviewed journal in Fall 2023.

Implications

The empirical knowledge revealed in this study would suggest policies and strategies to promote particular digital capabilities to develop resilience during the pandemic. They will serve as a guide for business managers and policy makers to effectively promote resilience during crisis situations in organizations and communities. They can motivate potential entrepreneurs to begin the digital entrepreneurial journey. For example, use of adaptive AIIS (Artificial Intelligence-based information systems) for planning and implementing non-pharmaceutical interventions (NPIs) such as lockdown of businesses can be effective to contain the pandemic (Baier et al., 2020).

We aim to provide insights and guidance to administrators of educational institutions to harness the power of sophisticated technologies that can be used to enable all learners to reach their full potential. For example, new 3-D technologies are capable of creating virtual learning environments that approach reality itself. Networking technologies are rapidly changing, affording low-latency, high-bandwidth solutions that can support learners almost everywhere. Artificial intelligence in the form of 24-hour mentors, research assistants and creativity stimulators can be used to design education platforms and enhance learning efficacy.

Conclusion

We believe that this empirical study addresses a two-fold gap in current literature: one, the critical role of pandemics in strategy and entrepreneurship (Agarwal et al., 2017; Chandra, 2018) and two, the pandemic as a stimulus to the evolution of digital entrepreneurship. In contemporary research digital entrepreneurship is viewed as the intersection of digital technology and entrepreneurship. We expect this research stream to serve as a stepping stone to inform researchers and practitioners by highlighting the characteristics of digital capabilities and consequences of a pandemic as drivers of digital entrepreneurship.

References

Agarwal, R., Moeen, M., & Shah, S. K. (2017). Athena's birth: Triggers, actors, and actions preceding industry inception. *Strategic Entrepreneurship Journal*, *11*(3), 287-305.

Alvarez, S. A., & Barney, J. B. (2007). Discovery and creation: Alternative theories of entrepreneurial action. *Strategic Entrepreneurship Journal*, *1*(1-2), 11-26.

Amit, R., & Zott, C. (2001). Value drivers in e-business. *Strategic Management Journal, 22*, 493-520.

Autio, E., Nambisan, S., Thomas, L. D., & Wright, M. (2018). Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems. *Strategic Entrepreneurship Journal*, *12*(1), 72-95.

Baier, Lucas; Schoeffer, Jakob; and Kühl, Niklas, "Utilizing Adaptive AI-based Information Systems to Analyze the Effectiveness of Policy Measures in the Fight of COVID-19" (2020). ICIS 2020 Proceedings. 4. https://aisel.aisnet.org/icis2020/is_health/is_health/4

Baker, T., & Nelson, R. E. (2005). Creating something from nothing: Resource construction through entrepreneurial bricolage. *Administrative Science Quarterly, 50*(3), 329-366.

Barreto, I. (2012). A behavioral theory of market expansion based on the opportunity prospects rule. *Organization Science, 23*(4), 1008-1023.

Brynjolfsson E., McAfee A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. New York: Norton.

CDC (2020). Retrieved from https://www.cdc.gov/surveillance/surveillance-data-strategies/Modernizing-for-Threats.html, accessed January 23, 2021.

Contact Tracing Apps. (2020). Retrieved from https://privacy.org.nz/assets/2020-05-12-OPC-Comparison-of-COVID-19-Apps-colours.pdf, accessed January 23, 2021.

Chandra, Y. (2018). Mapping the evolution of entrepreneurship as a field of research (1990–2013): A scientometric analysis. *PloS one*, 13(1), e0190228.

Davidsson, P. (2019). Guidepost: Look out! See change? Sea change ahead! *Academy of Management Discoveries.* doi.org/10.5465/amd.2019.0141.

Davidsson, P., Recker, J., & von Briel, F. (2018). External enablement of new venture creation: A framework. *Academy of Management Perspectives.* doi.org/10.5465/amp.2017.0163.

Gibson, J. J. (1977). The theory of affordances. In R. E. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing* (pp. 67–82). Hillsdale, NJ: Lawrence Erlbaum.

Gregoire, D., & Shepherd, D. (2012). Technology-market combinations and the identification of entrepreneurial opportunities: An investigation of the opportunity-individual nexus. *Academy of Management Journal*. *55*(4), 753-785.

Gudi, A., Chinta, R. (2020) "Awareness and Possession of Digital Competencies for Digital Entrepreneurship". AMCIS 2020 TREOs. 13. https://aisel.aisnet.org/treos_amcis2020/13

Hiatt, S. R., Sine, W. D., & Tolbert, P. S. (2009). From Pabst to Pepsi: The deinstitutionalization of social practices and the creation of entrepreneurial opportunities. *Administrative Science Quarterly,* 635-667.

Mooney, John; Ives, Blake; Agarwal, Ritu; Fitzgerald, Brian; Massey, Anne P.; Sambamurthy, Vallabh; and Soh, Christina, "Business school deans and institutional leader perspectives on the impacts and implications of the COVID-19 pandemic for the IS field and academic community" (2020). ICIS 2020 Proceedings. 3. https://aisel.aisnet.org/icis2020/panels/panels/3

Nadella, S. 2020 “Crisis requires co-ordinated digital response,” Financial Times, April 30, 2020 accessed at https://www.ft.com/content/b645d2f8-89f9-11ea-a109-483c62d17528

Nambisan, S. 2017. Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship. *Entrepreneurship Theory and Practice 41*(6), 1029-1055.

Porter, M. E., & Heppelmann, J. E. (2014). How smart, connected products are transforming competition. *Harvard Business Review, 92*(11), 64-88.

Rai, A. (2020). The COVID-19 Pandemic: Building Resilience with IS Research. *MIS Quarterly, 44*(2) iii-vii.

Contact Tracing Apps. (2020). Retrieved from https://privacy.org.nz/assets/2020-05-12-OPC-Comparison-of-COVID-19-Apps-colours.pdf, accessed January 23, 2021.

Ramoglou, S., & Tsang, E. W. (2016). A realist perspective of entrepreneurship: Opportunities as propensities. *Academy of Management Review*, *41*(3), 410-434.

Recker, J., and von Briel, F. (2019). "The Future of Digital Entrepreneurship Research: Existing and Emerging Opportunities" in *Fortieth International Conference on Information Systems, Munich 2019.*

Shane, S. (2003). *A General Theory of Entrepreneurship: The Individual-Opportunity Nexus*. Cheltenham, UK: Edward Elgar.

von Briel, F., Davidsson, P., & Recker, J. 2018. Digital technologies as external enablers of new venture creation in the IT hardware sector. *Entrepreneurship Theory and Practice, 42*(1), 47-69.

Williams, T., Gruber, D., Sutcliffe, K., Shepherd, D., Zhao, E. (2017). Organizational Response to Adversity: Fusing Crisis Management and Resilience Research Streams. *The Academy of Management Annals.*

Wood, M. S., & McKinley, W. (2010). The production of entrepreneurial opportunity: A constructivist perspective.

Review of Cloud Computing Research: A Pilot Study

Abstract

This paper is intended to start a research project reviewing recent studies related to various cloud computing issues. These issues include (1) cloud-related capabilities and agility, (2) impact of cloud on organizational and inter-organizational structure, (3) cloud promotion and marketing, (4) greening the cloud, (5) cloud security, and (6) cloud neutrality. These studies are either empirical or analytical in research methods. With the initial review of selected studies, potential gaps are identified with implications for future cloud research. Lastly, categorization of reviewed studies is also developed to guide future efforts in reviewing cloud related studies.

Introduction

As AWS, Google Cloud, and Microsoft Azure make more profits from their cloud services and more companies migrate to the cloud, many issues related to cloud computing are worthy of studying and data are also available for researchers to explore and generate insights. The intention of this paper is to initiate a research project reviewing recent cloud related studies to point out potential areas for future research with fruitful findings. In this paper, cloud related issues covered include (1) cloud-related capabilities and agility, (2) impact of cloud on organizational and inter-organizational structure, (3) cloud promotion and marketing, (4) greening the cloud, (5) cloud security, and (6) cloud neutrality.

A Pilot Review of Cloud Computing Related Studies

Cloud-related Capabilities and Agility

With survey data, Liu et al. (2018) developed and verified a model about the impact of cloud infrastructure (CI) flexibility and integration on organizational agility. The construct of organizational agility consists of operational agility (i.e., quick response to customer changing product and service needs), partnering agility (i.e., quickly setting up governance mechanisms and processes) and customer agility (i.e., cooperation with customers to explore and exploit opportunities). CI flexibility refers to the degree to which the cloud-based IT infrastructure can deal with volatility of service quality, fluctuations in service requirements, rapid business changes, scalability requirements, and new business relationships. CI integration refers to the degree to which the cloud-based IT infrastructure supports data access and integration of various applications both within a firm and across firm boundaries. Model test empirically shows significant and positive impacts of CI flexibility and integration on organizational agility. Further, it is also found that cloud computing spending positively moderate the impact of CI flexibility and integration on operational and customer agility. This paper contributes to the development of cloud-based IT competences and explores their impacts on agility, extending the resource-based view.

Kathuria, et al. (2018) empirically tested a model about how various cloud related capabilities may improve business flexibility (in terms of both responsiveness and scalability) and firm performance. They proposed cloud related capabilities including such as technological capability (CTC), service portfolio capability (CSPC) (consisting of service offering capability and market offering capability), and integration capability (CIC) (consisting of cloud legacy synchronization capability and cloud legacy consistency capability). Further, their model indicates that (1) CTO drives both CSPC and CIC, (2) CIC drives CSPC, (3) both CSPC and CIC drives business flexibility (BF) and firm performance, and (4) BF impacts firm performance positively. These cloud related capabilities and their impacts on business flexibility and firm performance supplements/complements findings in Liu et al. (2018).

With interview data, Krancher et al. (2018) developed a conceptual model regarding how elasticity and abstraction of the PaaS-based CICD support software development agility by providing both low and high levels of affordances. First, the affordance of shaping the environment (as the low-level affordance) in terms of the capacity of developers in setting up the infrastructure (with the PaaS-based CICD) to test their newly developed programs which triggers continuous feedback (as the high-level affordance) for fast learning leads to software development agility. Second, the abstraction feature of the PaaS-based CICD not only enables environment shaping for fast learning but also facilitates reuse of both software components (as another low-level affordance) for software product and development process innovations. The reuse of software components not only reduces efforts needed for continuous feedback, but also eliminates some dependencies on the infrastructure team by generating self-contained tasks which enables self-organization among developers to advance continuous feedback. Third, it is also identified in the model that both functional sub-teams and dependencies among different applications inhibit the formation of self-contained tasks. This conceptual model elucidates the rationales for the PaaS-based CICD in agile software development, demonstrating the value of cloud computing for software development in general.

Impact on Organizational and Inter-Organizational Structure

With migration to the cloud, at the organizational level, he changing role of the IT department has to be addressed. Among many role-related questions, the question of whether the IT department should be treated as a cost center, or a profit center is structure oriented. Based on the analytical model in Choudhary and Vithayathil (2013), it is recommended that if for some cloud services, the provider has a high pricing power, then the IT department is better to be treated as a profit center so that the firm may benefit from the structure as the IT department is motivated to customize the cloud services to add value to the consuming units in the firm in the sense that those units may make more profits. For some other cloud services, if the cloud provider does not have a high pricing power, then the IT department is better treated as a cost center so that all the consuming units may benefit from the free use of those cloud services. These results are based on

a supply chain model applied to the cloud provider – client relationship. The key is that the IT department as the value adding mediator between the cloud provider and the firm internal cloud service consumers could either charge those consumers (by deciding the quality and price of services) or provide free services to those consumers (while the firm decides the quality and bear costs of those services). While this paper provides some anecdotes, empirical tests are needed to verify those analytical results.

At the inter-organizational level, Manuel, et al. (2019) did a survey study and empirically identified the positive impact of cloud computing assimilation on supply chain integration. In their definition, supply chain integration consists of integration of financial flow, physical flow, and information flow. Cloud computing assimilation is about the level of cloud use in the organization and the level of cloud use consisting of micro cloud (for testing purpose), private cloud (for an individual firm), community cloud (for a group of related firms), and public cloud (for heterogeneous external organizations freely associated twitch each other through a wide global network). The significant impact of cloud assimilation on supply chain integration clearly indicates the impact of cloud computing on inter-organizational structure and process.

Cloud Service Promotion and Marketing

Shi (2022) explored the role of the AWS sponsored YouTube community (as an interactive knowledge management system) in potentially transforming community members into competent clients and empirically tested a model about the intertwined relationships among quality content development, community participation, and cloud provider's financial performance. Further, it is also indicated that when allocating resources in developing content and facilitating community participation, both the long-run survival and short-run success should be considered. In addition, to develop the conceptual model (about the relationships among content, participation, and value creation), the structure of the YouTube community AWS sponsors is conceptualized based on the well-tested onion model of the OSS community structure and it is indicated that there are three layers including junior enthusiastic learners, senior enthusiastic learners, and advanced clients. This paper demonstrates the role of social media in promoting and marketing cloud services in the sense of knowledge dissemination and customer relationship management.

Liu and Li (2019) explored the conditions under which various promotional programs may benefit the provider the best. They indicated that the provider normally decides the cloud service price and the promotion program used in a targeted or untargeted manner. Then the clients decide whether to use the services. Specifcially, through mathematical modeling, these authors investigated both coupons and free trials delivered through either targeted or untargeted manner under the condition of complete or incomplete information about the clients. With complete information about clients, they found that targeted programs are always better than untargeted programs and indicated that a high usage cost, a lower quantity of usage, and weaker network effects are reasons why targeted programs are better for the provider. If these conditions are not

met, then the targeted trial program is the most profitable promotional program. With incomplete information about clients, they found that no promotional program and untargeted coupons/trials generate the same profit. Lastly, they explored the situation where the client may not be aware of the valuations of the cloud services prior to adoption and found that trial programs are more advantageous due to the learning effects through trials. Empirical studies are needed to further verify these findings.

Greening the Cloud

As more people are using the Internet for various activities (such as ecommerce and social media) and organizations migrate their IT infrastructure into the public and private cloud, energy cost due to the increasing number of data centers around the world is also increasing dramatically over time. Kumar, et al. (2022) pointed out that one way to green the cloud is to balance the number of jobs on each resource server as much as possible. To achieve this goal, they developed a load balancing algorithm minimizing job waiting time in the cloud while also ensuring minimum QoS. They conducted a simulation based on the algorithm to demonstrate that an understanding of the optimal price for each job (with respect to waiting time) helps organizations decide the best time to transfer jobs to a public cloud from their private clouds (to reduce private data center costs)

Cloud Security

Cloud security consists of both the client-side (external-facing) security and the firm-side (internal facing) security. Username/password and second authentication are typical examples of client-side security and security measures (such as system monitoring and server logs) taken in the firm IT infrastructure in a cloud are typical examples of firm-side security. As the client side security level is increasing, the usability of the system from the client perspective likely decreases. With analytical models, Liu et al. (2015) explored the strategies taken by cloud vendors when facing competition in mixed or pure markets. Mixed market means that cloud vendors provide similar services (such as Microsoft Office 365 and Google Docs) with different objectives (such as profit maximizing and brand building). Pure market means that cloud vendors provide similar services for profit (such as Microsoft OneDrive and Dropbox). They found that in a mixed market, a profit-maximizing vendor always selects a higher level of client-side security than a brand-building firm and the total value surplus can be maximized when the brand building firm is the market leader. It is also indicated that the first mover advantage exists in this market. In a pure market, the optimal client-side security level is determined by the ratio between the cost of security breach and the user's fit cost (and this ratio is low if users value usability more than security). The fit cost is related to product customization capability which is addressed in Zhang (2020) (as will be reviewed next).

Zhang et al. (2020) developed a game-theoretical model to compare cloud-based services with on-premises software services in the context of low vs. high security-loss and probabilities of security attack for both types of systems (cloud vs. on-premise). Their model comprises two stages for both

vendors (cloud vs. on-premises). The first stage is about vendor decisions regarding security and product customization capability investment and the second decision is about the vendor service pricing. They had some interesting findings. For example, instead of the common thinking regarding the less secure cloud-based services, it is found that in a high security-loss environment, the expected security loss is less with cloud-based services (compared with on-premises based services). Further, it is also found that if the probability of attacks on cloud services is large enough, the expected loss due to cloud-based services is low in a high security-loss environment (compared with low security-loss environment). They also found that on-premises vendor's investments are dependent on consumer security loss. While in a low security-loss environment, an increase in customization investment by an on-premise vendor decreases its investment in security (i.e., they are substitutes), in a high security-loss environment, an increase in customization causes a greater investment in security for on-premise software systems. Again, while these cloud security related studies do generate insights about cloud pricing, security, and product customization, empirical studies are needed.

Cloud Neutrality

With the increasing impact of cloud computing on performance of various types of organizations differing in size, cloud capabilities, and financial resources, revenue for cloud providers and fairness for all cloud users may trade off with each other. Joe-Wong and Sen (2018) developed an analytical model (incorporating user demand, provider pricing, and cloud resource allocation decisions) to quantify how different users' cloud resource requirements may create a tradeoff between provider revenue and user fairness. They found that except that the user's service requirements are sufficiently symmetric or asymmetric, tradeoffs do exist. Their model may help regulators decide whether intervention is needed to achieve cloud neutrality from cloud resource allocation perspective.

Conclusion

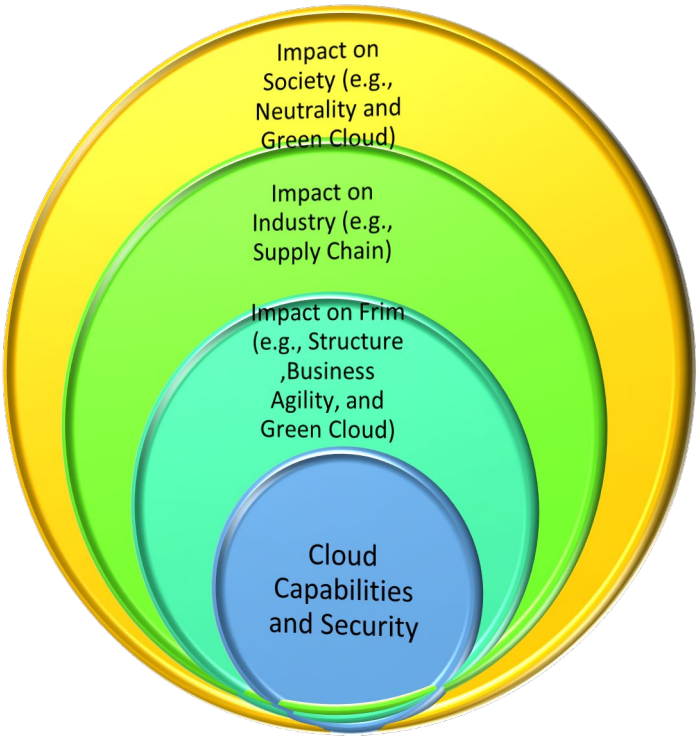
Table 1 in the following summarizes the reviewed studies with 2 dimensions including issues about cloud computing and research method applied. Both empirical and analytical methods are applied in extant cloud computing research related to impact of cloud computing on organizational and inter-organizational structure and cloud promotion/marketing. However, for cloud related capabilities and their impacts, analytical method-oriented study is missing. For issues about greening the cloud, cloud security, and cloud neutrality, empirical method-oriented studies are missing. These missing links could be due to the initial review being not comprehensive or they may also point out potential areas where additional research efforts could be made to advance our understanding of cloud related issues. Further, these cloud related research topics may be categorized into (1) cloud capabilities and security from both the provider and the client, (2) impact of cloud on firm (structure, business agility, and green cloud), (3) impact of cloud on industry (e.eg., supply chain), and (4) cloud neutrality and green cloud at the society level. Particularly, green cloud is related to both the firm level and the society level as reduced energy cost implies a

lower-level carbon dioxide emission. Figure 1 illustrate these different categories with the cloud capabilities and security at the core.

Table 1: Selected Studies related to Cloud Computing

	Research Method	
Issues with Cloud Computing	***Empirical***	***Analytical***
Cloud-related Capabilities and Business Agility	Liu et al. (2018); Kathuria, et al. (2018); Krancher et al. (2018)	
Impact on Organizational and Inter-Organizational Structure	Manuel, et al. (2019)	Choudhary and Vithayathil (2013)
Cloud Service Promotion and Marketing	Shi (2022)	Liu and Li (2019)
Greening the Cloud		Kumar, et al. (2022)
Cloud Security		Liu et al. (2015); Zhang et al. (2020)
Cloud Neutrality		Joe-Wong and Sen (2018)

Figure 1: Categories of Cloud Related Studies



References

1. Choudhary, V. and Vithayathil, J., 2013. The impact of cloud computing: Should the IT department be organized as a cost center or a profit center? *Journal of Management Information Systems*, 30(2), pp.67-100.
2. Joe-Wong, C. and Sen, S., 2018. Harnessing the power of the cloud: Revenue, fairness, and cloud neutrality. *Journal of Management Information Systems*, *35*(3), pp.813-836.
3. Kathuria, A., Mann, A., Khuntia, J., Saldanha, T.J. and Kauffman, R.J., 2018. A strategic value appropriation path for cloud computing. *Journal of management information systems*, *35*(3), pp.740-775.
4. Krancher, O., Luther, P. and Jost, M., 2018. Key affordances of platform-as-a-service: Self-organization and continuous feedback. *Journal of Management Information Systems*, 35(3), pp.776-812.
5. Kumar, C., Marston, S., Sen, R. and Narisetty, A., 2022. Greening the Cloud: A Load Balancing Mechanism to Optimize Cloud Computing Networks. *Journal of Management Information Systems*, *39*(2), pp.513-541.
6. Liu, S., Chan, F.T., Yang, J. and Niu, B., 2018. Understanding the effect of cloud computing on organizational agility: An empirical examination. *International Journal of Information Management*, *43*, pp.98-111.
7. Liu, Y. and Li, S., 2019. An Analysis of Promotional Programs for Cloud Computing: Coupons or Free Trials? *International Journal of Electronic Commerce*, *23*(3), pp.405-426.
8. Liu, Y., Sheng, X. and Marston, S.R., 2015. The impact of client-side security restrictions on the competition of cloud computing services. *International Journal of Electronic Commerce*, *19*(3), pp.90-117.
9. Manuel Maqueira, J., Moyano-Fuentes, J. and Bruque, S., 2019. Drivers and consequences of an innovative technology assimilation in the supply chain: cloud computing and supply chain integration. *International Journal of Production Research*, 57(7), pp.2083-2103.
10. Shi, Z., 2022. Relationship between social exchanges and financial performance: examining a virtual community sponsored by a cloud provider. *International Journal of Electronic Commerce*, *26*(3), pp.355-387.
11. Zhang, Z., Nan, G. and Tan, Y., 2020. Cloud services vs. on-premises software: Competition under security risk and product customization. *Information Systems Research*, 31(3), pp.848-864.

Decision Making: Public Administration and Policy

DSS, Machine Learning, and Artificial Intelligence

ADVANCEMENTS IN COMPUTATIONAL RHETORIC: A SURVEY

Wally Rodriguez
Villanova University, Villanova, PA, USA
wrodrig5@villanova.edu

Q B. Chung
Villanova University, Villanova, PA, USA
q.chung@villanova.edu

ABSTRACT

Computational rhetoric is a relatively new field of study with limited research output up until now. However, the topic is gaining interest both from the theory-building fronts and in development of tools. In this paper, we present a survey of progress in the field of computational rhetoric.

KEYWORDS: Computational rhetoric, linguistics, rhetorical argumentation, artificial intelligence, metaphor, genre analysis

INTRODUCTION

Computational tools are implemented techniques in computers to solve problems by either stepwise, repeated, or iterative solution methods. They use information technology to process large amounts of data algorithmically, as well as applied mathematics such as probability or statistical analysis. Computational tools also take advantage of such up-and-coming technologies as artificial intelligence and machine learning. New developments in these contributing fields add to the insights that computational tools can offer to existing disciplines.

Computation is generally regarded as a tool equipped with logical and deductive capabilities. Scientific disciplines have paved the way for the advancement of computational application, e.g., “computational physics” to automate heavy-duty calculations with precision. In the computational branches of various established disciplines, use of computational tools is on the rise in general research activities. However, disciplines considered more qualitative in nature, such as history, law, and linguistics, also have enjoyed benefits from developing computational applications. The use of computational tools produces a “qualitatively different scientific approach” that gives rise to a new domain of questions in such fields (Wojcik, 2013).

Although the field of linguistics has enjoyed a robust development of its branch of computational linguistics, the same may not be said of an important subfield of linguistics, rhetoric. Rhetoric may be defined as the art of persuasion, or more specifically as the use of figures of speech and compositional techniques. In addition, it may be defined as the “study of how ideas can be shaped and shuffled to fit together

and be coaxed into fitting onto the structure of existing ideas (Wojcik, 2013). When compared to linguistics, which studies the theory of language structure, the field of rhetoric is more focused on a higher-level analysis of language having to do with features such as sensitivity to the audience, understanding of psychological features and points of view of speakers and listeners, subjectivity of truth and acceptability, and context dependence (Crosswhite, et al., 2003). These aspects of rhetoric may be altogether recognized as rhetorical theory. The argument for the development of computational rhetoric is that “scale is computation’s main bounty” (Graham, 2020).

COMPUTATIONAL RHETORIC GAINING MOMENTUM

The field of computational rhetoric, when only considering strictly the field of computational rhetoric and not related fields, is fairly limited in regard to research work done on it. There are a few research groups working on the field of computational rhetoric- such as a group of researchers from the University of Waterloo (Computational Rhetoric Workshop, 2016). However, the topic is gaining more interest as publications on computational rhetoric have been included in academic journals. (See, for instance, Majdik, 2019). Moreover, there now exists a journal devoted to the study of techniques for computational rhetoric, *Argument & Computation* (IOS Press, n.d.).

Implementation of computational tools for the study of linguistics is becoming increasingly relevant as applications of speech and text become a part of everyday life. One such application may be the widespread use of natural language in search engines, as well as smart assistant devices that respond to natural languages. In addition, there are applications geared to the professional knowledge work. The following are pursuits that have grown out of computational linguistics with computational-rhetoric applications:

> Text summarization: This is the process of breaking down lengthy text into digestible paragraphs or sentences. This method extracts vital information while also preserving the meaning of the text. This reduces the time required for grasping lengthy pieces such as articles without losing vital information. For example, multiple news articles could undergo a text summarization software, producing a digestible summary. This is a helpful tool when wanting to derive insights from many long strings of text.
>
> Argument mining: This is the automatic identification and extraction of the structure of inference and reasoning expressed as arguments presented in natural language. Argument mining solutions are used in the field of law to automatically extract arguments and reasoning from legal documents (Atkinson & Bench-Capon, 2021).
>
> Sentiment analysis: This is arguably the best-known technique of natural language processing to extract subjective information from a text such as microblog posts and customer reviews. It is used to determine the attitude and polarity of a document. Another application may be detection of sentiments

> toward companies' earnings calls and relevant corporate documents to assess the optimism or lack thereof of the leadership of a company.

Rather than a pure computational linguistic focus on modeling discourse, computational rhetoric takes into account rhetorical argument discourse (Grasso, 2002). In the development of computational rhetoric models, discourse theories can examine the structure of text to identify the interactions between words, adverbs, verbs, and other parts of speech. Rhetorical structure theory defines rhetorical relations that can hold between spans of a text (Walton & Gordon, 2017).

A computational rhetoric tool would take into account argumentation based on evaluative judgments, audience, how arguments are perceived, linguistic style of the presentation of the argument to the audience (Grasso, 2019). Computational tools serve to extend the field of rhetoric by enabling rhetorical projects that are impossible without them. When taking into account natural language processing systems, it is important to incorporate the recognition of rhetorical speech. In addition, rhetorical figuration is linguistically and cognitively inescapable since it exists in all textual content, ranging from speeches to novels to social media posts.

ADVANCEMENT OF RESEARCH IN COMPUTATIONAL RHETORIC

Some research has been conducted on the incorporation of rhetorical frameworks into computational tools. Incorporating rhetorical effectiveness into computational tools is important to obtain better results when using automated approaches to analyzing speech. These tools require an interdisciplinary approach including natural language processing, discourse analysis, argumentation theory, and computational models of argumentation.

Here is a sample of recent advances in computational rhetoric research.

MetaNet

A large team of researchers from ICSI, UC San Diego, University of Southern California, Stanford, and UC Merced is building a computer system capable of understanding metaphors used in multiple languages. According to them:

> "The MetaNet project seeks to systematically identify and analyze the metaphors that people use to discuss and reason about a broad range of topics and domains. To accomplish these aims, we have developed a system that makes use of a repository of formalized frames and metaphors to automatically detect, categorize, and analyze expressions of metaphor in large-scale text corpora." (MetaNet, n.d.)

Metaphors are an important component of rhetoric. By focusing on the identification of metaphors, this project is making an important advancement for computational rhetoric tools.

RhetFig Project

The RhetFig Project, or the Rhetorical Figure Ontology Project, represents an important advancement in computational-rhetoric research for computationally identifying rhetorical figures. The project reports the following:

> "The Rhetorical Figure Ontology Project, a multidisciplinary research project that is working towards the development of a comprehensive database of rhetorical figures, an associated wiki, and ultimately, an ontology of rhetorical figures. The database and wiki project provide the dataset and space for the conceptual development, respectively, to create an ontology." (Kelly, et al., 2010).

Statistical Genre Analysis

When analyzing bodies of text, it is important to determine the genre of text when gathering insights from it. Guided by rhetorical genre theory, statistical genre analysis attempts a computational method to identify the predicted patterns of recurring moves that are associated with stable genres of writing. In this case, this study created a coding schema inspired in statistical genre analysis to analyze a large number of transcripts from meetings from authorities in Health (Graham, 2020).

In doing so, it employs statistical genre analysis a methodology that combines techniques from rhetorical studies, linguistics, and health communication. SGA is our attempt to hybridize the humanistic and qualitative study of discourse with statistical methods. It is a method designed to move technical communication toward the ability to offer encompassing conclusions about large data sets in gathering insights on rhetorical inquiry. This type of study shows the power of computational-rhetorical analysis in demonstrating how textual features can result in certain outcomes across a "big data" set.

IBM Watson Debater

This is an artificial intelligence system designed and developed by IBM to debate humans on complex topics. This software is now being used as a decision support system by companies and enterprises.

> "Project Debater can use a curated corpus or crowd-sourced set of opinions to perform a key point analysis and narrative generation on a particular topic. This system has capabilities to support computational argumentation, fundamental natural language processing, and the organization and generation of content." (IBM Research, 2016).

Its AI capabilities represent an important development for computational rhetoric tools, given that it serves as a demonstration that computers can harness big data to engage in speechwriting, debate and argumentation.

APPLICATIONS OF COMPUTATIONAL RHETORIC TOOLS

Given the development of computational tools based on language and rhetoric frameworks, researchers on computational rhetoric have been able to utilize computational rhetoric tools for real-world purposes. As described below, these studies range from the detection of hateful or harmful rhetoric, to understanding the use of a key word through time by analyzing a large data set of texts.

HateBERT: Retraining BERT for Abusive Language Detection in English

In a study involving shorter spans of text, a retrained BERT model, a transformer-based machine learning technique for natural language processing developed by Google, was used to detect abusive language. The BERT model was retrained by using a large-scale dataset of Reddit comments from banned Reddit communities for being offensive, abusive, or hateful. This study represents an automated approach to understand and detect certain rhetoric (Caselli, et al., 2021).

Agentic Framing of Climate Change in the Congressional Record, 1994-2016

Identifying patterns of word usage in discourse, computational rhetoric tools can be used to evaluate the rhetorical effects as the propagation of rhetorical expressions across large sets of texts. In this study, computational rhetoric tools are employed to analyze how Congress has framed human agency toward addressing climate change.

This test case focuses on how the term of climate change is rhetorically used in Congressional deliberations over a 22-year time span by studying the rhetorical frame of human capability relative to climate change. He stated that computationally tracking how distinct rhetorical expressions increase or decrease in prevalence over time provides a framework for engaging questions about rhetorical effectiveness- for “mapping the life” of rhetorical expressions (Majdik, 2019).

The Opioid Epidemic and the Pursuit of Moral Medicine: A Computational-Rhetorical Analysis

A similar study employs computational-rhetoric analysis to study biomedical writing on opioids. It evaluated the framing of opioid pharmacology, the relative attention paid to pain management versus opioid dependence risks, and the distribution of statements related to physicians’ primary ethical obligations. In doing so, the study found a trend of changing bioethical responsibilities from individuals to public health. The employment of a computational-rhetoric model meant that the study appropriately contextualized the framing of the word “opioid” in each of its sample texts to derive its insights on the matter (Graham, 2020).

CONCLUSION

The development of computational rhetoric techniques is important in gaining a more contextual and better-framed analysis of narratives using computational language tools. Due to the relative novelty of the topic, researchers worldwide are still engaging in developing computational tools pertaining to rhetorical theories. It is a daring and complex task to automate the detection of rhetoric in text, given the subjectivity and variability of nuance of natural language. As computational tools and related technologies continue to develop, such as AI and machine learning, many of these research threads may converge and produce more effective output that can be more universally used.

REFERENCES

Atkinson, K. & Bench-Capon, T. (2021). Argumentation Schemes in AI and Law. *Argument & Computation*, 12(3), 417-434.

Caselli, T., Basile, V., Mitrović, J. & Granitzer, M. (2021). HateBERT: Retraining BERT for Abusive Language Detection in English. *Proceedings of the 5th Workshop on Online Abuse and Harms (WOAH 2021)*. DOI: 10.18653/v1/2021.woah-1.3, 17-25.

Computational Rhetoric Workshop. (2016). Computing Figures, Figuring Computers. 18 May 2016. https://computationalrhetoricworkshop.uwaterloo.ca/. (Accessed on 12/28/2022).

Crosswhite, J., Fox, J., Reed, C., Scaltsas, T., Stumpf, S. (2003). Computational Models of Rhetorical Argument. In: Reed, C., Norman, T.J. (eds) *Argumentation Machines. Argumentation Library*, vol.9. Springer, Dordrecht.

Graham, S. (2020). The Opioid Epidemic and the Pursuit of Moral Medicine: A Computational-Rhetorical Analysis. *Written Communication*, 38(1), 3-30.

Grasso, F. (2002). Towards Computational Rhetoric. *Informal Logic*, 22(3), 195-229.

Grasso, F. (2019). A Mental Model for a Rhetorical Arguer. *Proceedings of EuroCogSci 03*, 139–144.

IBM Research. (2016). Project Debater. https://researcher.watson.ibm.com/researcher/view_group.php?id=10774. (Accessed on 12/28/2022).

IOS Press, (n.d.), Argument & Computation, https://www.iospress.com/catalog/journals/argument-computation. (Accessed on 12/28/2022).

Kelly, A. R., Abbott, N. A., Harris, R. A., DiMarco, C. & Cheriton, D. R. (2010). Toward an Ontology of Rhetorical Figures. *Proceedings of the 28th ACM International Conference on Design of Communication*, 123-130.

Majdik, Z. P. (2019). A Computational Approach to Assessing Rhetorical Effectiveness: Agentic Framing of Climate Change in the Congressional Record, 1994–2016. *Technical Communication Quarterly*, 28(3), 207-222.

MetaNet. (n.d.). The MetaNet Wiki. https://metanet.icsi.berkeley.edu/metanet/. (Accessed on 12/28/2022).

RhetFig. (n.d.). https://artsresearch.uwaterloo.ca/chiastic/display/index.php. (Accessed on 12/28/2022).

Walton, D. & Gordon, T. F. (2017). How Computational Tools Can Help Rhetoric and Informal Logic with Argument Invention. *Argumentation*, 33(2), 269-295.

Wojcik, M. W. (2013). Inventing Computational Rhetoric. Master of Arts Thesis. Michigan State University.

Visualizing DOE Results with Classification Trees: A Supply Chain Application

Jiaoxue (Michelle) Liu, jxliu98@bu.edu
John Maleyeff, maleyeff@bu.edu
Adrian Perez, adrianbu@bu.edu

Boston University
Metropolitan College

Abstract

The application of analytical tools to support decision making in business has benefitted from visualizations that provide intuitive meaning. The emergence of big data analytics has led to an increase in analytical sophistication that has made multi-factor decision making powerful but less intuitive. This article explores the integration of two methodologies: (a) more traditional experimental design and (b) more recent machine learning. Specifically, data visualization using a classification tree is used to create a decision tool derived from data collected using a central composite experimental design. This approach combines data collection efficiency with improved decision-making effectiveness. An example application stemming from a project with a business-to-business manufacturer who needed to determine if new products should be make-to-stock or make-to-order is used as illustration. Results suggest that, although some loss of precision occurs, a classification tree provides promising results that were well received by supply chain managers at the client firm.

1. Introduction

Many data scientists consider themselves storytellers because they strive to make sense of information or data sets often derived from complex processes. The methods they employ use a variety of visualizations that help decision makers obtain an accurate, comprehensive, yet intuitive understanding of important implications. Classification trees can be very useful because they provide an easy-to-follow interactive decision procedure. They are consistent with how many decision makers approach challenges using informal or heuristic algorithms. For example, a medical professional may diagnose a patient's illness based on a series of questions such as: Does the patient have a fever above 38 degrees? If yes, does the patient have a sore throat? If yes, the patient most likely has a cold (not the flu).

This article continues a recent trend among researchers to use machine learning visualizations in conjunction with more traditional design of experiment (DOE) methodologies. DOE has been a mainstay of industrial statistics for many years because it creates cost-efficient yet powerful experimental designs. The intent of this work is to use classification trees to analyze a DOE experiment. It is intended to enhance a decision makers ability to apply the DOE results. The approach was motivated by a recent project whereby a client asked for a set of simple decision rules after a model was created to determine if a new product should be make-to-order or make-to-stock.

The article is organized as follows. Brief introductions are presented to DOE and classification trees, and recent efforts to integrate these methodologies are described. This section is followed by the presentation of a robust methodology using a central composite experimental design. An example application, based on the project that motivated this work, is described from both technical and decision-making perspectives. A comparison of the tradition

DOE analysis approach and the application of a classification tree methodology follows. The various classification tree level options are evaluated to determine the most useful structure, whose accuracy is then analyzed.

2. Background

In the early 20th century, Ronald Fisher introduced a new approach that revolutionized experimental analysis and led to the development of statistical methods known as the design of experiments (Fisher, 1971). George E.P. Box played a significant role in adapting Fisher's methods for industrial applications, including the oft-quoted mantra "Every process generates information that can be used to improve it" (Box, 1987). The use of DOE for non-traditional uses includes Taguchi's inner-outer arrays for creating robust designs (Bryne and Taguchi, 1987). Many textbooks are available that detail the use of DOE in practice, including the use of various experimental designs such as randomized blocks, factorial, and Placket-Burman (e.g., Box et al, 2005).

In a DOE experiment, one or more factors are purposefully set at two or more levels to analyze their effect on an experimental response, which is typically a performance metric or other outcome of interest. Various experimental structures exist that enhance experimental efficiency based on an assumed underlying model that quantifies the response based on the factor levels. For example, a 2^k factorial design consists of k factors that are set at 2 levels each; this design assumed a linear model. A DOE experiment is typically analyzed using either analysis of variance or regression to determine the factors that significantly impact the response. The graphical results for the significant factors are presented as a main or interaction plot that shows how the response is affected by changes in the significant factors (Montgomery, 2019, p. 194).

Data visualization has a long history, with early examples of pictographs (i.e., picture-writing systems) discovered about 3100 BCE to record events or tell stories (Friendly & Wainer, 2021, p. 11). The evolution of visualizations began to broaden the modes of presentation in various forms, including five primary groups: geometric, pixel-oriented, icon-based, hierarchical and graph-based techniques (Keim & Kriegel, 1996). For instance, the bar chart first appeared in 1786 to illustrate imports and exports in trade (Playfair, 1786), and histograms followed to demonstrate compression and approximate data distributions (Ioannidis, 2003). Economy attributes, such as stock market patterns, have been displayed using a tree map (Wattenberg, 1999). Finally, linear discriminant analysis projects multi-dimensional data onto a one-dimensional space (Leban et al, 2006).

Recent developments in machine learning, big data, and data mining have brought renewed interest to the development of visualizations that create intuitive displays to help decision makers make sense of statistical results. Visualizations have been a mainstay of quality methods in supply chains (e.g., Ishikawa, 1982), and many other industrial applications. They can require creativity to intuitively display multiple dimensions simultaneously (Tufte, 2001). Today, visualizations have become more difficult because data are more expansive, and analysts often lack intuitive insights prior to their analysis (EIU, 2014). Visualizations can be subjective, and care must be taken to present objective, yet intuitive, displays (Van Wijk, 2005). When successful, decision makers will make better and meaningful choices in a logical manner (Bose & Mahapatra, 2001). They have been applied to healthcare (Park et al, 2022), logistics (Gürdür et al., 2018), and medicine (Gotz & Borland, 2016), to name a few.

Integrating DOE with machine learning and other newer methodologies is underway. Arboretti et al (2021) provide a review of literature pertaining to the use of machine learning and DOE for product innovation; they focus on the use of DOE as a means for presenting active learning scenarios for the machine learning algorithms. These methods are shown to be applicable for product innovation in the chemical industry (Arboretti et al, 2022). Freisleben et al (2019) propose the integration of DOE with various machine learning methods for quality improvement; they conclude that the use of DOE can increase the speed at which machine learning results are obtained. Hamad et al (2022) point to improvements in understanding factor interactions when DOE is integrated with machine learning.

This article concerns the use of classification decision trees to show results of purposefully designed experiments. Classification decision trees, or more simply classification trees, disaggregate heterogeneous information into a set of homogeneous groups (Camm et al, 2016, p. 456). Lau et al (2015) used classification trees to determine optimal levels for feature selection based on a structured experiment with an iterative search approach. Their application uses classification trees to inform their experimental design, while the application presented here uses DOE to provide data to the classification tree algorithm.

3. Methodology

The central composite design (CCD) is an experimental design that augments a 2^k factorial experiment with additional center-point experiments and a set of axial points (Box et al, 2005, p.450). It includes a combination of experiments whereby each factor ($x_i, i = 1,2,\cdots,k$) is set at -1 and $+1$ levels, plus one or more center point experiments (level 0) and a set of axial experiments (level α) as shown in Equation 1. The CCD generates a second-order response surface as shown in Equation 2, where y is the response and the residual variation (ϵ) is assumed to be normal.

$$\alpha = 2^{k/4} \quad [1]$$

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{j=2}^{k} \sum_{i<j} \beta_{ij} x_i x_j + \epsilon \quad [2]$$

For the example application detailed in Section 4, a CCD was developed consisting of $k = 4$ factors, as shown in Table 1, where each row corresponds to an experimental combination that will generate one or more responses. The (0,0,0,0) center point is typically repeated multiple times so that residual variation can be quantified.

Development of a classification tree requires an algorithm that creates a group of homogeneous clusters from a multi-dimensional data set. In the supply chain application below, a classification tree is created using the data obtained from the CCD experimental design. The methodology uses a mean square error calculation to find the optimal set of groupings where the user can indicate the number of levels in the classification tree. The intention of the classification tree is to provide a robust yet intuitive mechanism for a decision maker.

Table 1: CCD with $k = 4$

x_1	x_2	x_3	x_4
-1	-1	-1	-1
1	-1	-1	-1
-1	1	-1	-1
1	1	-1	-1
-1	-1	1	-1
1	-1	1	-1
-1	1	1	-1
1	1	1	-1
-1	-1	-1	1
1	-1	-1	1
-1	1	-1	1
1	1	-1	1
-1	-1	1	1
1	-1	1	1
-1	1	1	1
1	1	1	1
-2	0	0	0
2	0	0	0
0	-2	0	0
0	2	0	0
0	0	-2	0
0	0	2	0
0	0	0	-2
0	0	0	2
0	0	0	0

4. Example Application

A consulting project was undertaken for a manufacturer of plumbing, heating, and other water-related products in the Winter of 2022. The problem statement concerned their logistics and distribution network that supplies customers (i.e., retailers) located in every region of the US. The firm operated several domestic and international manufacturing facilities, and it leased warehouse space in five strategically located distribution centers (DC's). Some products are available to customers from the nearest DC, which holds safety stock and quotes a one-week lead time to deliver those products to customers. Alternatively, some products are make-to-order from a manufacturing facility. In these cases, lead times are longer, and the firm holds no product inventory. Transportation costs differ, with more frequent and smaller shipments in the make-to-order scenario.

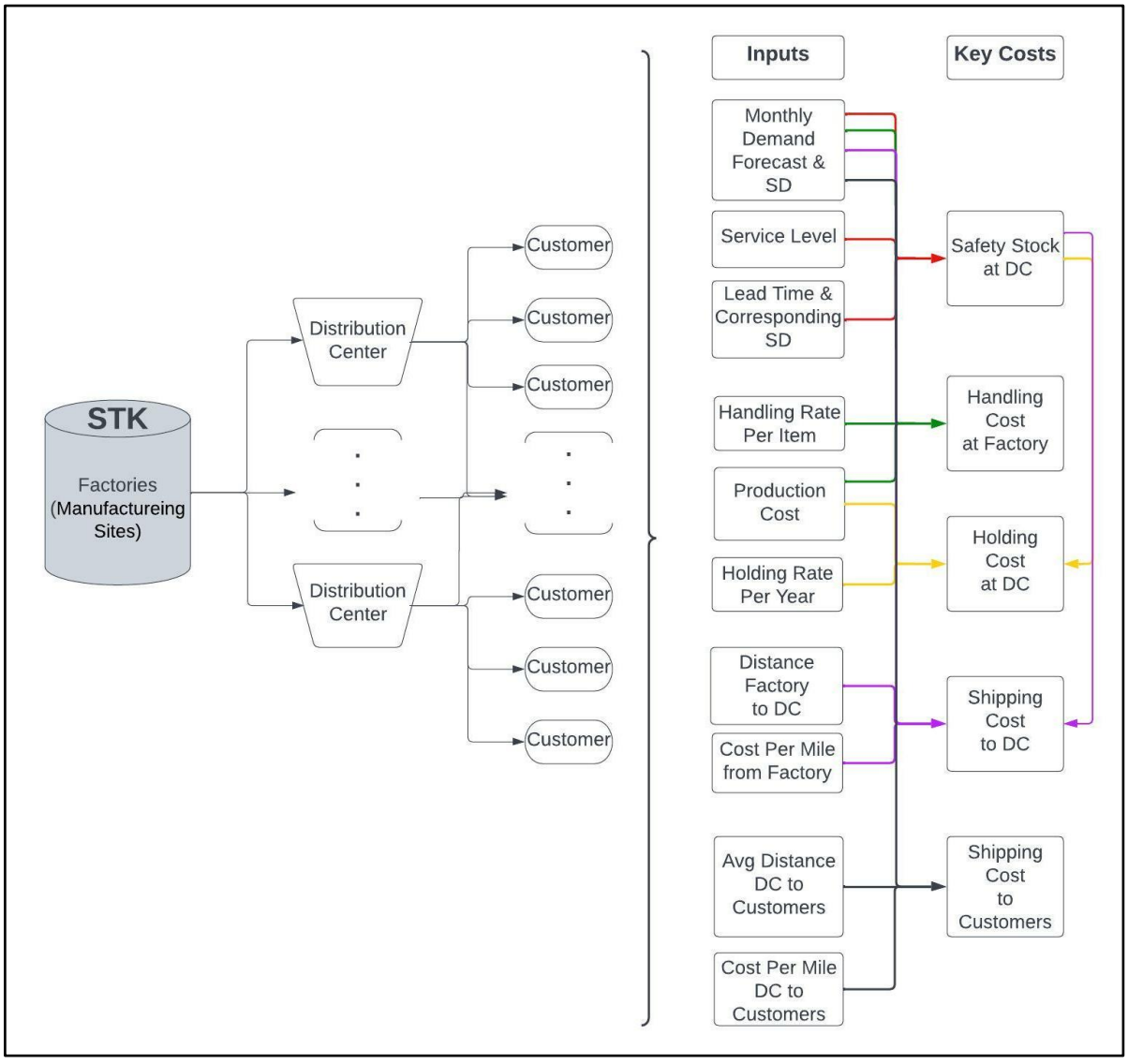
The project team created a comprehensive inventory-distribution model to guide decisions (i.e., make-to-stock vs. make-to-order) for new products. Inputs included product and distribution parameters, costs, and forecast accuracy estimations. Outputs included the annual total cost for each alternative. This analytical model satisfied the project's technical requirements; however, the client requested a set of relatively simple rules for making future decisions. This desire stemmed from their discomfort with a relatively complex decision tool where, at times, precise input data were unavailable. Although the use of DOE in this application does not concern experimentation with random outcomes, DOE is employed whereby the responses are derived from the analytical model and used to evaluate the efficacy of the proposed method.

Figures 1 and 2 illustrate the make-to-stock (STK) and make-to-order (MTO) options, including the process flow and annual cost model framework. The difference in annual costs (called *delta* with label d), determines the outcome of the decision model (Equation 3).

$$d = \frac{Y_{MTO} - Y_{STK}}{Y_{STK}} \quad [3]$$

The sign of d determines the decision. When d is positive, the STK option is less expensive than the MTO option (and vice-versa).

Figure 1: Make-to-Stock Option (STK)

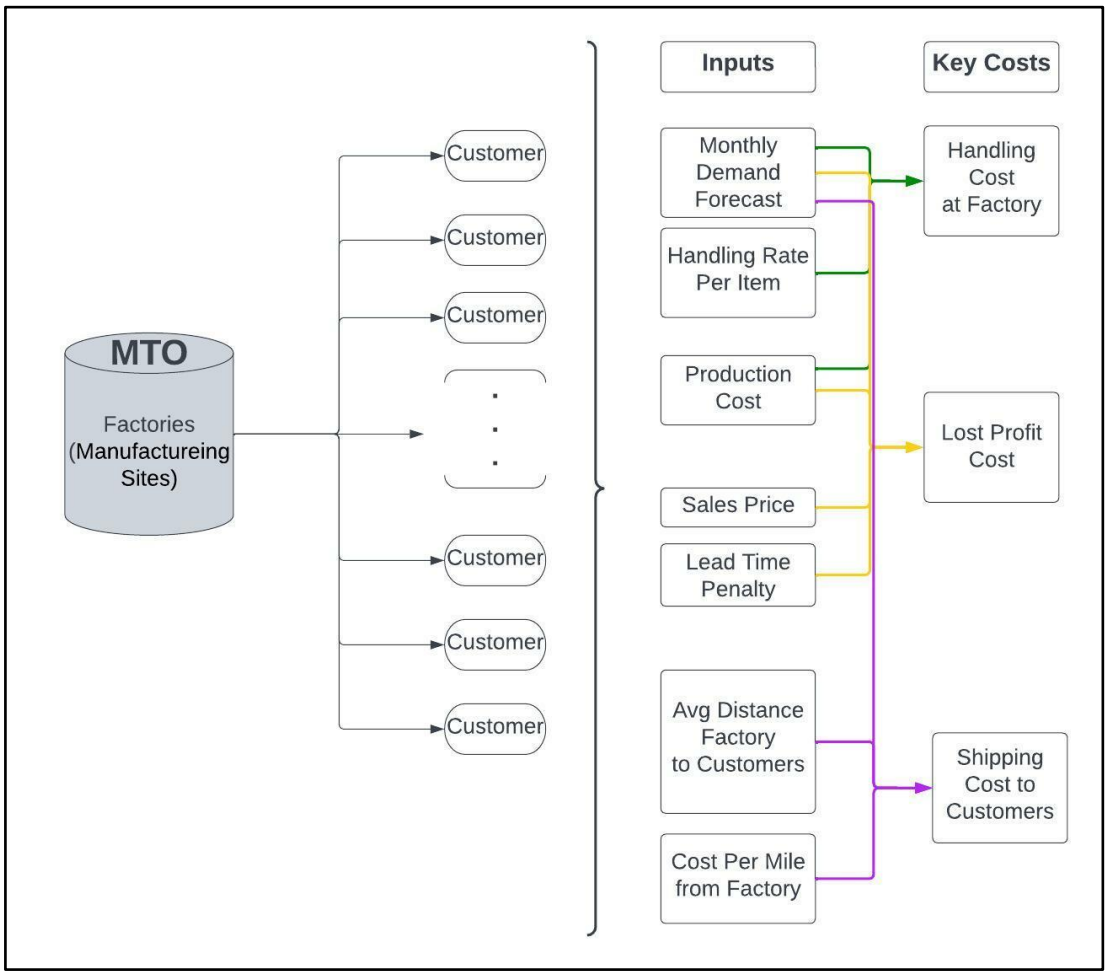


A preliminary analysis identified four factors as being crucial to the decision. A CCD was created using these factors, which are listed below (including the definition of their levels):

- Production cost (C) per unit (-1 = $30, +1 = $70).
- Demand forecast (V) coefficient of variation (-1 = 35%, +1 = 65%).

- Lead time (L), factory to DC or factory to customers (-1 = 52 days, +1 = 116 days).
- Average shipping distance in thousands of miles (D), factory to DC or factory to customers (-1 = 1 thousand miles, +1 = 2 thousand miles).

Figure 2: Make-to-Order Option (MTO)



4.1 DOE Results

The CCD in Table 1 was run and a regression analysis was used to determine significant factors and to create a mathematical model. Because the researchers already knew that the four factors and almost their interactions were statistically significant, the focus was on the creation of the mathematical model and the visualization of results using interaction plots. The second-order model, derived from the CCD experiment, is shown as Equation 4.

$$\hat{d} = 1.1383 + 0.0101x_C + 0.0069x_L - 4.0780x_V - 0.3x_D - 0.00003922x_C^2 + 0.00001071x_L^2 + 3.0571x_V^2 + 0.02309x_D^2 + 0.00002249x_Cx_L - 0.0119x_Cx_V + 0.0005838x_Cx_D - 0.0091x_Lx_V - 0.0007443x_Lx_D + 0.4x_Vx_D \quad [4]$$

DOE main and interaction plots show the relationship between factor levels (for significant effects) and the response. Two interaction plots for the DOE results are shown as illustration.

Figure 3 indicates that the value of d is lower as the forecast CV increases and d is lower for lower production costs especially when the CV is low.

Figure 3: Interaction Plot: Forecast CV and Production Cost

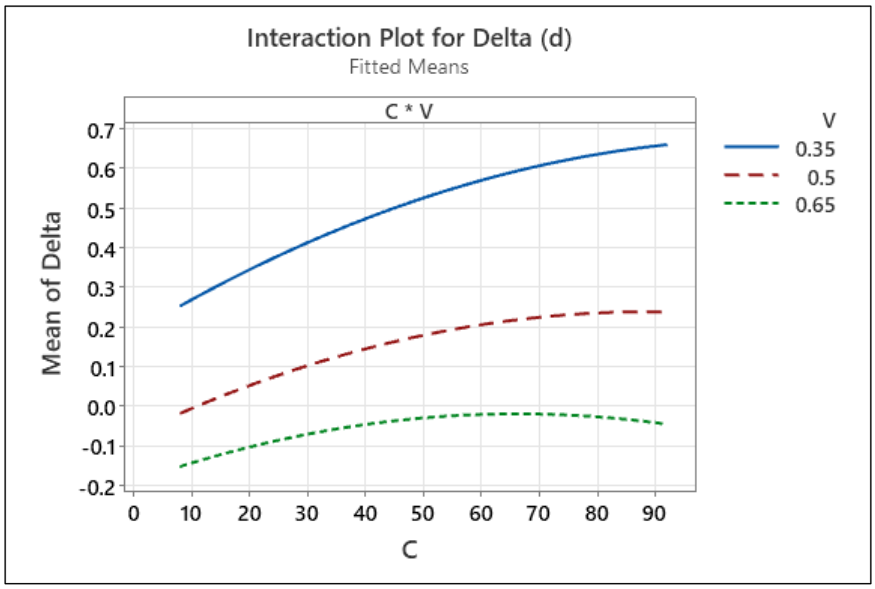
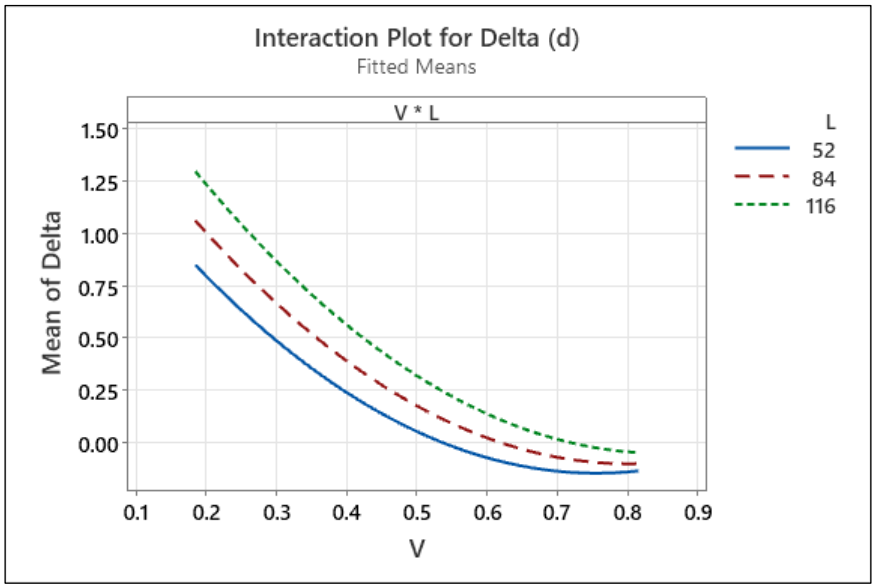


Figure 4 indicates that the value of d is lower as the forecast CV increases and that d is especially lower for shorter lead times (L).

Figure 4: Interaction Plot: Lead Time and Forecast CV



To a decision maker, the insights provided by the interaction plots are interesting but fall short of providing them with impactful decision criteria. For example, Figure 3 implies that higher production costs will favor the make-to-stock option, but only when the demand CV is low. When the demand CV is high, the production cost has little effect on the decision. The supply chain managers at the client firm found these and other implications interesting but

unhelpful. Their focus was not on the effects, but on what combination of factor levels should cause them to choose one option over the other. Besides, interaction plots are difficult to make use of when more than two factors combine to impact a decision. Based on these observations and input from the client, the research team decided to explore the use of a classification tree.

4.2 Classification Tree Results

Three classification trees of varying depths were created form the data generated by the CCD. They were evaluated for both their accuracy and usefulness. In the Appendix, Figures A1-A3 show the 3-level, 4-level, and 5-level classification trees, respectively. The following examples are presented that illustrate the use of classification tree with depth 4 from a user's perspective:

a. If the new product has a forecast CV less than 42.5%, the decision should be make-to stock when the lead time is less than 68 days.
b. If the new product has a forecast CV greater than 42.5%, the decision should be make-to-stock when the lead time is longer than 132 days.
c. If the new product has a forecast CV between 42.5% and 72.5%, the decision should be make-to-stock when the lead time is between 68 and 132 days.

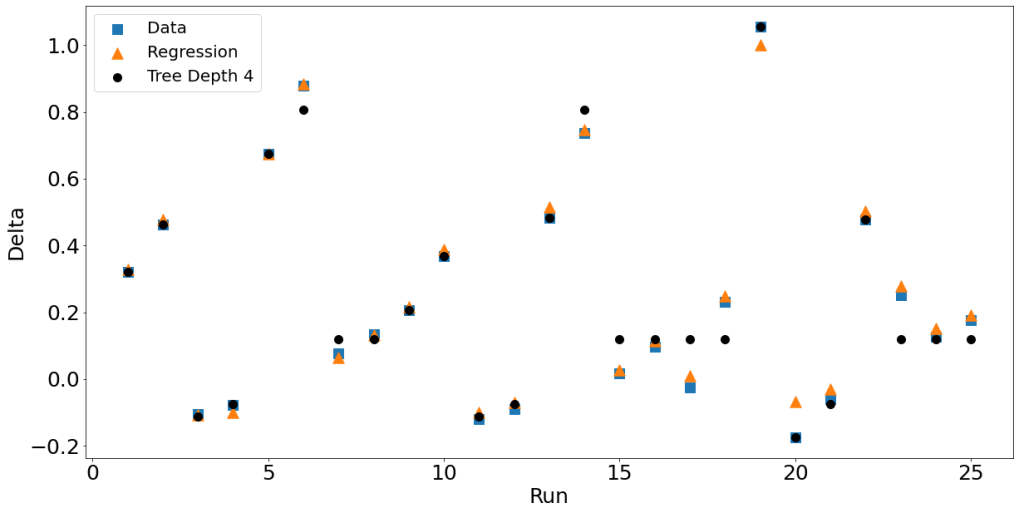
The classification tree can be presented to clients as a binary choice (MTO or STK), or it can include both the binary decision and the continuous cost comparison that adds additional context. For example, in the 4-level classification tree we can see that the make-to-stock option is especially beneficial for new products have a forecast CV less than 42.5% and lead times greater than 68 days.

4.3 Accuracy Comparison

Although prediction power improves as the classification tree depth increases, there is a decrease in the rate of improvement as more levels are added. Here, the overall accuracy of the 3-level tree was 82% and the overall accuracy of the 4-level and 5-level trees were both about 88%. Higher-level trees are also less useful for decision makers because they fail to achieve their intent of simplifying decision making. The 3-level tree, although simpler in application, was much less accurate than the 4-level tree. The 5-level tree was almost identical in accuracy with the 4-level tree, and it was considered too complex from the decision maker's perspective. Based on a comparison of these alternatives, the 4-level classification tree was selected as the focus of the accuracy comparison detailed below.

The two comparisons included an evaluation of the predictions associated with the experimental data and a simulation using random inputs parameters. The existence of the analytical model enabled the simulation analysis to be performed. The scatterplot shown as Figure 5 illustrates the relationship of the two prediction models with the CCD data. It is noted that the classification tree's strength (i.e., simplification) is also a source of inaccuracy because it predicts a finite set of potential response predictions. The simulation consisted of 1000 random sets of inputs based on a uniform distribution within the range of CCD settings. Predicted delta (d) values were converted into either the MTO or STK binary category, so that the user's perspective was maintained. A confusion matrix was used to compare the two approaches, with two metrics: (a) accuracy when the analytical model indicated STK and (b) accuracy when the analytical model indicated MTO.

Figure 5: Comparison of Accuracies (4-Level Tree)



This comparison is powerful because there is an underlying analytical model that can be considered perfectly accurate, and the two analysis approaches can be compared to this model. The comparison takes the decision maker's perspective. It assumes that when the difference between the two options (i.e., d) is less than 7.5%, either make-to-stock or make-to-order can be an effective decision. Table 2 shows the accuracy of the approaches as compared to the analytical model (called Actual in the table). The regression model is somewhat more accurate (it agreed with the actual MTO result 98% of the time versus 93% of the time for the classification tree). In addition, the regression approach agreed with the actual STK result 100% of the time versus 96% of the time for the classification tree. This difference in accuracy is offset by the improved effectiveness of the classification tree approach.

Table 2: Comparison of Accuracies

			Actual	
			MTO	STK
PREDICT	MTO	4-Level Tree	93%	7%
		Regression	98%	2%
	STK	4-Level Tree	4%	96%
		Regression	0%	100%

5. Conclusions & Future Work

The results of this comparison should be considered preliminary because they are based on a specific example, where most of the outcomes favored the make-to-stock option. Although the client for the project that motivated this work was happy with the classification tree approach, this example was not able to evaluate the use of classification trees in DOE in a more robust manner. The researchers plan to get feedback from fellow experimenters and data scientists, so that their approach can be improved. Going forward, a comprehensive set of example applications will be used to generate additional comparisons.

6. References

Arboretti, R., Ceccato, R., Pegoraro, L., & Salmaso, L. (2021). Design of Experiments and machine learning for product innovation: A systematic literature review. *Quality and Reliability Engineering, 38,* 1131-1156.

Arboretti, R., et al (2022). Machine learning and design of experiments with an application to product innovation in the chemical industry. *Journal of Applied Statistics, 49*(10), 2674-2699.

Bose, I. & Mahapatra, R. K. (2001). Business data mining: A machine learning perspective. *Information & Management, 39*(3), 211-225.

Box, G.E.P. and Bisgaard, S. (1987). The Scientific Context of Quality Improvement, *Quality Progress, 20*(6), 54-61.

Box, G.E.P., Hunter, J.S., & Hunter, W.G. (2005). *Statistics for Experimenters 2/E*, John Wiley & Sons, New York.

Bryne, D. and Taguchi, S. (1987). The Taguchi approach to parameter design. *Quality Progress, 20*(12), 19-26.

Camm, J.D. et al (2016). *Essentials of Business Analytics*, Cengage Learning, Boston, Massachusetts.

EIU (2014). The future of business supply chains: How companies are using data and analytics to make their supply chains more connected and agile. *The Economist Intelligence Unit*. http://www.eiu.com.

Fisher, R.A. (1971). *The Design of Experiments 9/E,* Macmillan, New York.

Freiesleben, J., Keim, J., & Grutsch, M. (2020). Machine learning and design of experiments: Alternative approaches or complementary methodologies for quality improvement? *Quality and Reliability Engineering International, 36*(6), 1837-1848.

Friendly, M. & Wainer, H. (2021). *A History of Data Visualization and Graphic Communication*. Cambridge, MA and London, England: Harvard University Press.

Gotz, D., & Borland, D. (2016). Data-driven healthcare: Challenges and opportunities for interactive visualization. *IEEE Computer Graphics and Applications, 36*(3), 90-96.

Gürdür, D., Raizer, K., & El-Khoury, J. (2018). Data visualization support for complex logistics operations and cyber-physical systems. In 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Funchal–Madeira, Portugal.

Hamad, K., Obaid, L., Haridy, S., Zeiada, W., & Al-Khateeb, G. (2022). Factorial design–machine learning approach for predicting incident duration. *Computer-Aided Civil and Infrastructure Engineering*, 1-21.

Ioannidis, Y. (2003). The history of histograms (abridged). In Freytag et al, Proceedings 2003 VLDB Conference, 19-30.

Ishikawa, K. (1982). *Guide to Quality Control 2/E*, Asian Productivity Organization, New York.

Keim, D.A. & Kriegel, H. (1996). Visualization techniques for mining large databases: A comparison. Transactions on Knowledge and Data Engineering. *Special Issue on Data Mining, 8*(6), 923–938.

Lau, J.K., Nguyen, M.H., Liang, W., and Tan, K.K. (2015). Application of design of experiments to feature selection in ventilation tube applicator. IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society, 004277-004282.

Leban, G., Zupan, B., Vidmar, G., & Bratko, I. (2006). Vizrank: Data visualization guided by machine learning. *Data Mining and Knowledge Discovery, 13*(2), 119-136.

Montgomery, D.C. (2019). *Design and Analysis of Experiments 10/E*, John Wiley & Sons, New York.

Park, S., Bekemeier, B., Flaxman, A., & Schultz, M. (2022). Impact of data visualization on decision-making and its implications for public health practice: A systematic literature review. *Informatics for Health and Social Care, 47*(2), 175-193.

Playfair, W. (1786). *The Commercial and Political Atlas: Representing, by Means of Stained Copper-plate Charts, the Exports, Imports, and General Trade of England, the National Debt, and Other Public Accounts*. Debrett and Robinson, London.

Tufte, E.R. (2001). *The Visual Display of Quantitative Information 2/E*, Graphics Press LLC, Cheshire, Connecticut, USA.

Van Wijk, J.J. (2005). The value of visualization. In *VIS 05 IEEE Visualization,* 79-86. Institute of Electrical and electronics Engineers.

Wattenberg, M. (1999). Visualizing the stock market. In CHI'99 Extended Abstracts on Human Factors in Computing Systems, 188-189, ACM Digital Library.

Appendix

Figure A1: Classification Tree (Depth = 3)

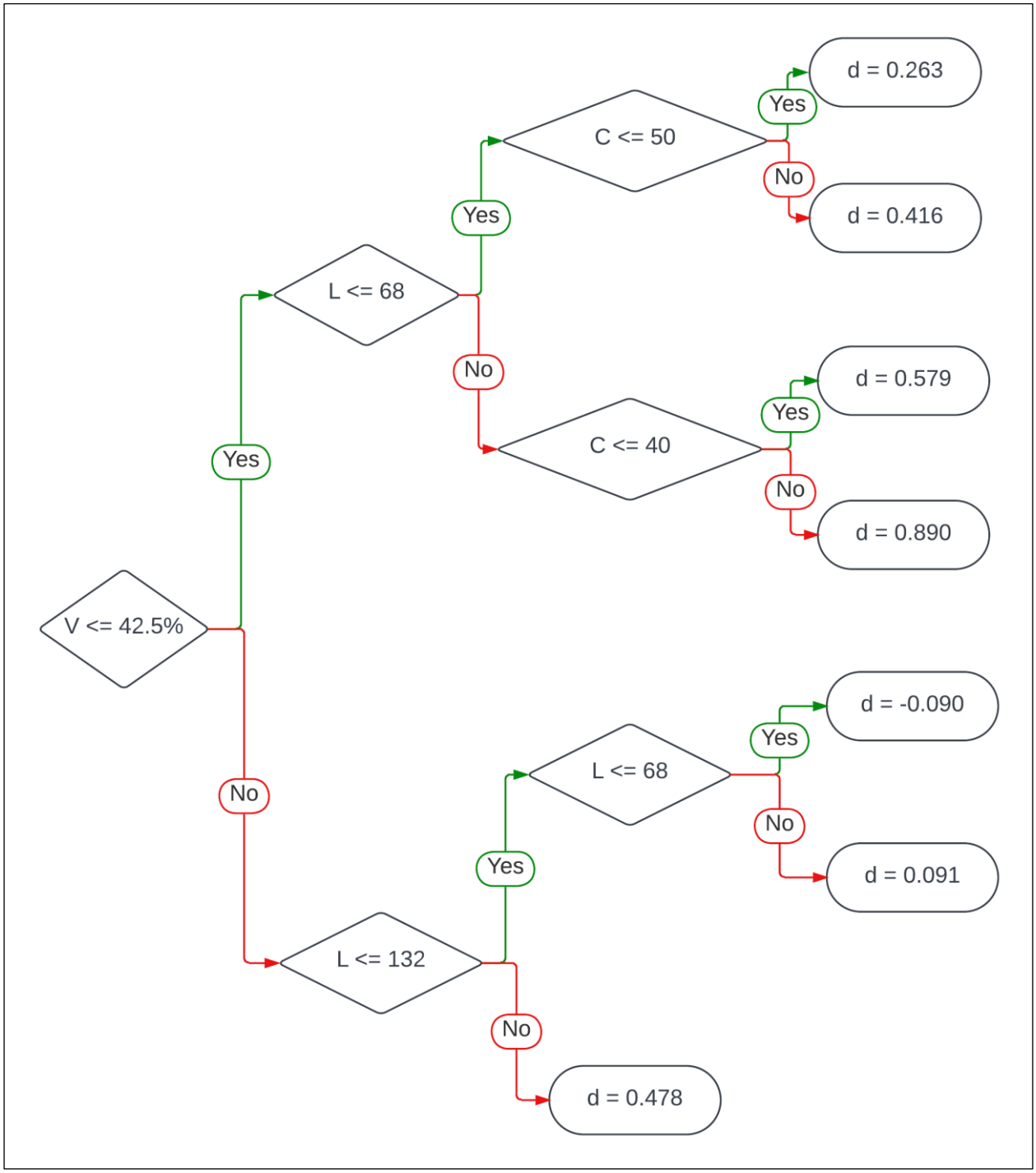


Figure A2: Classification Tree (Depth = 4)

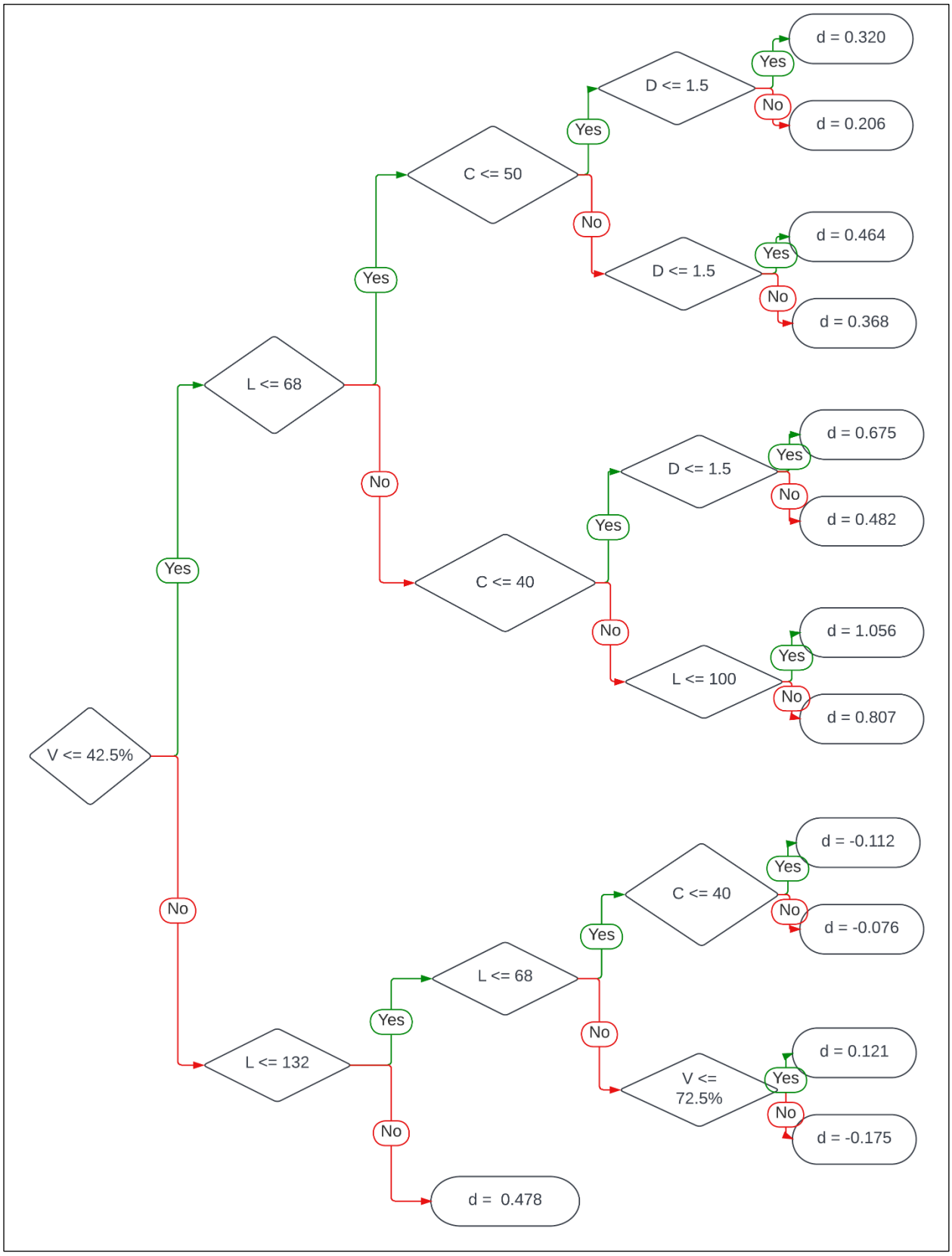
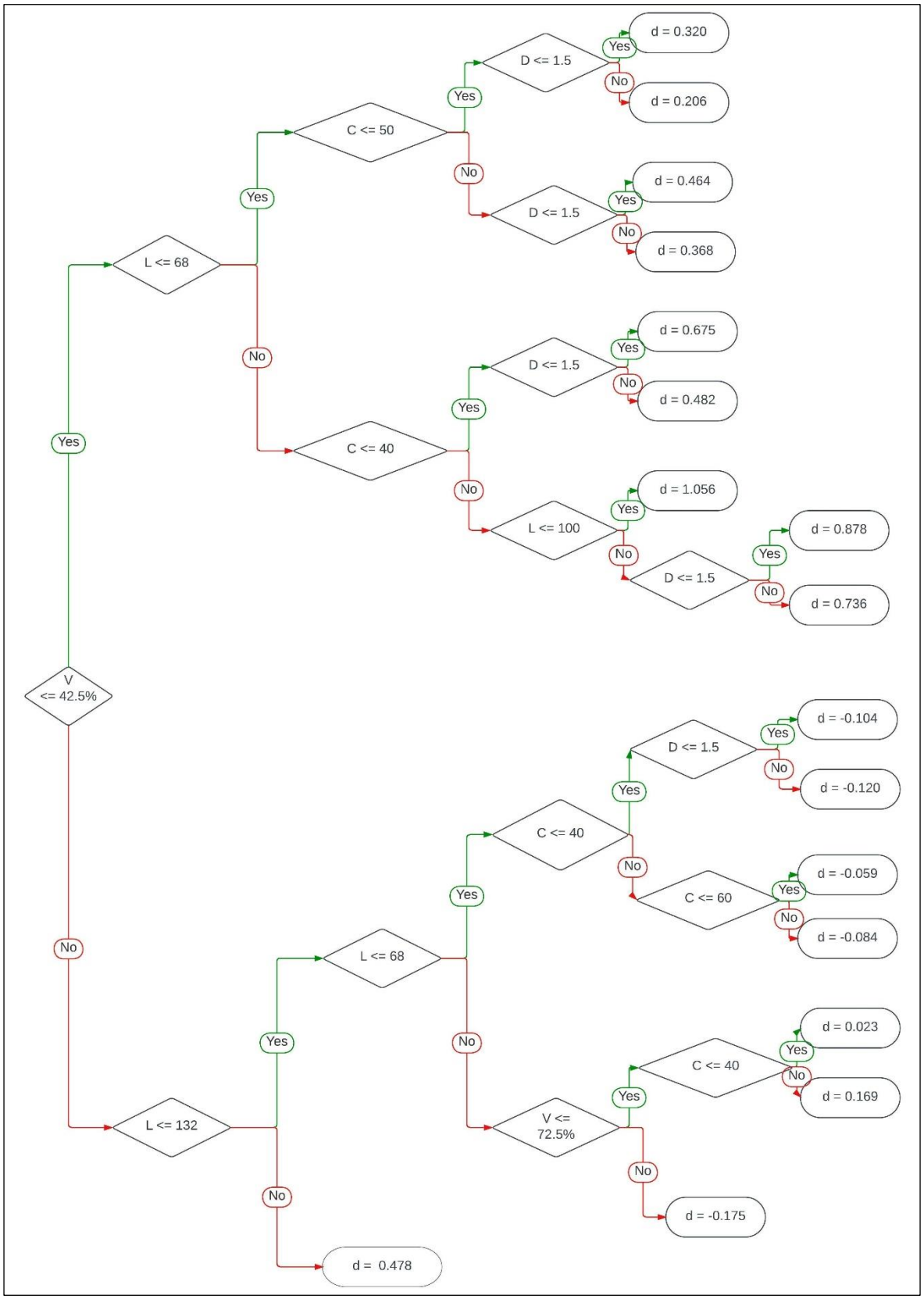


Figure A3: Classification Tree (Depth = 5)



Education, Curriculum, and Cases

Advancing Mathematical Modeling in K-12 Education

Thomas Davenport, in a noted 2006 article (Davenport, 2006), begins by reviewing a few of the problems companies must solve on a frequent – if not daily – basis, and lists a few companies noted at the time for their success in solving these problems. He then continues, stating

> …analytics competitors do all those things in a coordinated way, as part of an overarching strategy championed by top leadership and pushed down to decision makers at every level. Employees hired for their expertise with numbers or trained to recognize their importance are armed with the best evidence and the best quantitative tools. As a result, they make the best decisions: big and small, every day, over and over and over.

In the 17 years since Davenport penned his article, the amount, availability, and importance of data has continued to explode. Correspondingly, any field that involves analyzing and making decisions based on data has exploded.

Interestingly enough, 2006 was also the year I earned a graduate degree in the field of Decision Science. The methods and tools in decision science are used in a wide array of other fields: analytics, supply chain management, operations, marketing, finance, …to name a few.

What follows is the story of my journey into my present field, Decision Science, how that journey involved the development of my spatial abilities, how those spatial abilities came to – in part – define my life, and how – similar to Temple Grandin, a researcher with noted spatial abilities – my spatial abilities have both helped me in some areas and hindered me in other areas. This paper also serves to add another personal experience to concerns raised by researchers that young people with noted spatial abilities but weaknesses in either language or mathematical abilities – or both, could – with appropriate education, training, and mentoring – positively impact the significant shortages of qualified individuals in STEM/STEAM fields, both in industry and K-12 education. This paper is not a research summary but rather a combination of personal experience and position paper.

Literature Review

Temple Grandin wrote recently that her academic struggles, which were related to her enhanced spatial abilities, played a role in preventing her from pursuing her first choice of careers (Grandin, 2023). But she persevered and has become very well known in another – if related – field. The demand for individuals in STEM fields "…is expected to grow by almost 11% by

2031, over two times faster than the total for all occupations." (Krutsch and Roderick, 2022). In my mind, a strong link exists between these pieces of information. In fact, Lakin and Wai (2020, p. 1016) wrote expressly about this phenomenon:

> Although the societal value of spatially talented individuals is clear, traditional curricula continues to focus heavily on verbal reasoning and mathematical communication. Those students who excel based on their spatial skills do so despite a lack of traditional talent development opportunities found in most schools or curricula.

There are so many quotes from Professor Grandin's writings and Lakin and Wai's paper that I feel are relevant to me and could easily be included in this paper. More on that later.

Parallel to the calls in the literature for primary and secondary education improve the approach to harnessing and developing the abilities of students with spatial abilities, there is work in the mathematics education literature on developing students' modeling and application skills in mathematics. Kaiser and Maab (2007, p. 108) found that "…teachers and their beliefs about mathematics must be regarded as essential reasons for the low realization of applications and modelling in mathematics teaching." In the 15 years since that was written, progress has certainly been made. Students with spatial abilities may have a role to play in addressing this identified shortage.

My Journey

I don't think I was labeled an exceptional child. Maybe I should have been. This is my story.

In elementary school, math came easily to me and I enjoyed it. I was considered good at math. In fact, when I was in fifth grade, the teacher allowed me to work at my own pace through the textbook, instead of progressing at the pace of the rest of the class. I finished working through the textbook in March, if I recall correctly. I do not remember how I spent my time in math class for the rest of the year. Ironically, it was also in fifth grade that I began participating and competing in organized sports, which played a significant role in my life from that point until several years after I graduated from college.

In junior high school (7th grade, I think), I remember taking a standardized test that I thoroughly enjoyed taking. The test required the student to – in their mind – fold pieces of paper along dotted lines on the figure into three-dimensional objects. I don't remember if I ever learned how I scored on this test, but I am confident I scored well. It wasn't until a number of years into adulthood that I learned that this specific type of test has a name: object assembly test.

My freshman year of high school I began participating in a sport that was new to me – the sport in which I would eventually achieve the most success. In my sophomore year, a new coach arrived who was a higher caliber than what would be expected for the competitive level of our team. This coach – although I have not been in contact with him for many years – has had a significant impact on my life. Why? On a regular basis starting when he took over the program and for the duration of our time with him, he had us practice the motions of our sport with our eyes closed. My sport is one where form, although not a scored component of competition, plays a role in an athlete's success. It is generally acknowledged in the sport that – all things being equal – better form results in better athletic performance. This is a characteristic of many sports, and a characteristic on which has been placed increasing importance since I began participating in this sport. While I was practicing the motions of my sport with my eyes closed, I was able to visualize my body going through the motions as if I were in a helicopter (now a drone) above. I still do not know if any other athletes experience this phenomenon. In any case, and in part due to my ability to visualize myself performing the motions of my sport, the movie "Hoosiers" could have been written about our coach and our high school team. Partly as a result of this, I gained admission to and graduated from a competitive university.

In college, I did not distinguish myself academically. But sophomore year I took several architecture classes, enjoyed them while earning decent grades, and was asked to consider majoring in the field. I jumped at the chance. Unsurprisingly, the visual aspect of the field drew me in. Along with my architecture studies, I competed in intercollegiate athletics. I did not train extraordinarily hard, like some athletes who experience success do, but I achieved some measure of success. My athletic success in college, I believe, can be explained by two things:

1. The team's development of a culture of success over time
2. My athletic form, significantly enhanced by my ability to visualize myself performing the motions of the sport.

Overall, my college experience was pleasant although unremarkable. My visualization abilities were satisfied both by my architecture studies and my participation in athletics.

After college, I chose not to pursue architecture as a career. For the first several years I focused on training in my sport, finding a job that supported my athletic habit. My job happened to be in manufacturing, and I did not distinguish myself. After I ceased my athletic training, I happened to change jobs into a new industry. Like many industries over the last three decades, it is an industry where data has come to play a dominant role. Several years into this new-to-me industry, I began spending an increasing portion of my time working on the business side of information systems development and teaching myself and building office automation solutions. Early on I did not distinguish myself working with information systems. But towards the end of my tenure with the company a programmer in the systems development department that supported my division remarked that – at a recent department meeting – the systems development professionals in her systems development department had taken an informal poll and I was named “favorite business client” or the equivalent. Apparently the two characteristics that caused programmers to welcome the opportunity to work with me were that I knew what I wanted and I was willing to put in the work to achieve it. My visual abilities, again, played a role in my success in this area. That success came from my ability to visualize how information systems with which I am extremely familiar work, as I visualize transactions as they move through information systems. I do not know if others who work with information systems develop this ability. Over the course of my tenure with this company, however, there were a handful of occasions when I made decisions that – in retrospect – would not be made by an employee wishing to have a long and successful career with this, or maybe any, company. I made these decisions because I did not know better and did not think to ask for advice.

After leaving this company I entered graduate school and eventually earned two masters. As previously mentioned, one is in the field of Decision Science. After completing graduate school, I embarked on a career teaching as an instructor at the college level. For the past 15+ years, I have taught (overwhelmingly undergraduate) courses in statistics, operations, analytics, and information systems. To date I have not distinguished myself in my teaching career.

In 2017, after an unsuccessful attempt two years earlier, I was elected to my local school board. To say that serving on school board – especially since the COVID-19 pandemic started – has

been a learning experience would be a tremendous understatement. Our community is an incredibly diverse, high-poverty community. These characteristics are mirrored in our district's student population. Early in my school board service, programs and services in the areas of Social Emotional Learning (SEL) and Trauma-Informed Education (TIE) were introduced in our district. Emphasis on these initiatives – especially since the onset of COVID-19 – has increased in the district and statewide. I believe these initiatives provide significant benefits to our students and families, as they have so much with which to cope. While these initiatives help reduce, eliminate, or at least help manage, barriers to educational progress, I believe our district's offerings could be enhanced. I believe, as does Professor Grandin (2023), that schools should be places where students with visual abilities can be introduced to, and gain experience with, opportunities to put those visual abilities to use in ways that society views as positive. The alternative might be, as reported by some researchers including Lakin and Wai (2020), that students who have unmet needs related to their visual abilities are more likely to struggle academically and experience behavior issues. As these two researchers state (p. 1032), "This combined evidence demonstrates the importance of finding and serving spatially talented students."

One of Lakin and Wai's recommendations (2020, p. 1034-1035) is that "…teacher training and familiarity regarding spatial talents would help serve this neglected pool of students." Lakin and Wai continue, referring to Wai et al. (2009), that this is especially true because education majors don't, as a group, distinguish themselves in the area of spatial abilities. I completely agree with raising the awareness of spatial abilities in schools of education. I would like to propose that secondary students with spatial abilities have the opportunity to – while being provided with the appropriate guidance, support, and mentoring - become K-12 teachers themselves. Lakin and Wai (2020) speak of the possibility for those with spatial abilities to impact the STEM pipeline and middle skill job opportunities. I believe that those with spatial abilities can find success and be impactful in K-12 teaching, as I have found that teaching can have a creative element in the planning and delivering of lessons and preparation of materials. As I envision it, there might be one teacher per grade or several per school with spatial abilities. These teachers with spatial abilities could, through proper support and mentoring, be successful teachers in their own right and simultaneously be the creativity go-to people for their peers. And I personally believe that these teachers with spatial abilities would be ideally suited to serving as primary and secondary

math teachers using an approach to primary and secondary math education similar to that implemented at the tertiary level by and described by Pelesko, Cai, and Rossi (2013). Based on personal experience, I believe that such an appropriate to mathematics education is particularly well suited to students with spatial abilities.

Several final thoughts:

1. once in graduate school and twice in my post-graduate-school career there were brief episodes where my behavior was considerably outside of normal expectations.
2. As a school director I have not always followed the "go along to get along" approach.

These behaviors have impacted both me and those around me and played a significant role in landing me where I find myself today.

Conclusion

In conclusion, I believe that both the literature and my experience support the position that Professor Grandin has expressed over many years – and recently succinctly described as: "Society is failing visual thinkers, and that hurts us all." This paper is an attempt to link the literature to my experience and possible visual abilities, and also proposes an initiative that may advance mathematical modeling in K-12 education, help address the shortage of STEM/STEAM talent, and better serve individuals similar to myself.

REFERENCES

Davenport, T.H. (2006). Competing on Analytics. *Harvard Business Review*, *84*(1), 98-107

Grandin, Temple. "Temple Grandin: Society is Failing Visual Thinkers, and that Hurts Us All." *New York Times*, January 9, 2023, www.nytimes.com/2023/01/09/opinion/temple-grandin-visual-thinking-autism.html.

Kaiser, G. and Maab, K. (2007). Modelling in lower secondary mathematics classroom: Problems and opportunities. In W. Blum et al. (eds.), *Modelling and Applications in Mathematical Education*, The 14th ICMI Study, (pp. 99-108). New York: Springer Science+Business Media LLC.

Krutsch, E. and Roderick, V. (2022). STEM Day: Explore Growing Careers, *U.S. Department of Labor Blog, November 4, 2022*. https://blog.dol.gov/2022/11/04/stem-day-explore-growing-careers.

Lakin, J.M. and Wai, J. (2020). Spatially gifted, academically inconvenienced: Spatially talented students experience less academic engagement and more behavioral issues than other talented students. *Brit J of Edu Psychol*, *90* (1015-1038). DOI:10.1111/bjep.12343.

Pelesko, J.A., Cai, J., and Rossi, L.F. (2013). Modeling Modeling: Developing Habits of Mathematical Minds. In A. Damlamian et al. (eds.), *Educational Interfaces between Mathematics and Industry*, New ICMI Study Series 16, (pp. 237-245)). DOI: 10.1007/978-3-319-02270-3_23.

Does Management Education Need a Facelift? The Intersection of Managing, Leading, and Coaching. *Part III*

Maureen L. Mackenzie-Ruppel, PHD, PHR, CAPM, ICF-ACC
Full Professor of Business
Molloy University School of Business
Rockville Centre, New York 11571
(631) 682-0399
MaureenMackenzieRuppel@gmail.com

ABSTRACT
This paper continues the multi-phase exploration of management education, the growing profession of executive coaching, and its value to workplace preparation. It is the third in a series of five papers dedicated to this exploration. Part III summarizes the knowledge presented in both Part I and Part II. Part I included a literature review and discussion as to who defines premier business education, the growing profession of executive coaching, and higher education's responsibility to relevant workforce preparation. It set the stage for the multi-phase research plan focused on three sets of stakeholders: senior leaders/employers, business faculty, and MBA graduates. Part II reported insights from eleven in-depth interviews with senior leader as the future employers of our business graduates. It revealed what employers expect of business educators, as they prepare students to become their future leaders. **Part III** reports on the insights of eleven business faculty as relates to their role as the gatekeepers of business curriculum and student learning outcomes. Without the faculty, the integration of coaching skills, knowledge, and disposition into business education, will not be successful. The goal of this *multi*-paper exploration is to seek answers related to the question; *Does Management Education need a facelift?* If so, *should the coaching body of knowledge, skills, and theories, be integrated into management and business education*?

Keywords: Pedagogy, Coaching, Leadership, Manager Development.

INTRODUCTION TO THIS MULTI-PAPER SERIES OF SCHOLARLY PAPERS

A series of scholarly papers will explore management education and *coaching* as a meaningful body of knowledge, skills, and theories, for business managers and leaders. The discipline of coaching lives outside of the traditional business curriculum; at the most, it lives at its fringe. Yet, the skills and knowledge of coaching can serve senior leaders as they face their daily challenges. Coaching aligns with, and supports, participative leadership and empowerment. It serves leaders as they gain insights on their own thinking; it also serves them as they learn how to motivate and empower their followers. (DiGirolamo & Tkach, 2020; DiGirolamo & Tkach, 2019). The profession of executive coaching has expanded. Executive coaches guide and support valued talent to rise-up and be ready for the leadership opportunities that require self-awareness and thought-leadership. Yet, these leaders were not introduced to coaching as a legitimate management practice within their business curriculum and education (Filipkowski, Heverin, Ruth, 2019; Jamison, 2018).

As scholars and educators, we must explore whether coaching knowledge and skills should be actively integrated into the program-level learning outcomes for graduate business education? Currently, training for coach practitioners and internal executive coaches emerges from external organizations, rather than from the universities and colleges. Are business schools lagging behind this shift in leadership preparation?

WHY THIS RESEARCH IS IMPORTANT–REVIEW OF PART I (For full paper, see Mackenzie-Ruppel, 2021)

The Relevance of Higher Education to Student Success. In the March 6, 2020, *Chronicle of Higher Education*, ***prior*** to the shutdown of face-to-face education across the United States, it was reported that 30% of colleges face some market risk with 10% of colleges facing *severe* market risk. At the 2020 Northeast Business & Economics Association Conference, the influencing factors on higher education under-performance *prior to the 2020 Pandemic*, were presented (Mackenzie-Ruppel, 2020). They included Public Doubt and the questionable confidence that existed in the value of a college education. The student debt crisis

and the high cost of college led students and parents to demand tangible economic outcomes for the investment (The Trends Report, 2019, CHE).

Premier Workforce Preparation. Though business schools should take the lead as the premier source of business education, there is a threat of professional organizations becoming the '*go-to*' educational providers, as is the case with *executive coaching*. Also, the Mega-university offers a degree at a low cost with high convenience that allows students to earn the degree with ease and focus on industry certifications to prepare them for the workplace (Mackenzie-Ruppel, 2020).

Business Education Defined. There is *no* universal set of requirements that define *good* education for a discipline. Education oversight emerges from four viewpoints: (1) state government holds the authority to permit an institution to grant a degree (Program Registration, 2020), (2) academic disciplines and their voluntary national associations, (3) educational institutions and their regional and national associations, and (4) the federal government with its reach from the Department of Education with funding for students and colleges with regional accreditation (Harcleroad, 1980, Hegji, 2018, Harcleroad & Eaton, 2005).

Business Education Accreditation. The Association to Advance Collegiate Schools of Business (AACSB) connects educators, students, and business so these stakeholders can develop the curriculum for the next generation of leaders; yet there is no explicit clarity on curriculum *content*. Recently, there has been an intentional move away from a model or prescribed curriculum, leaving the institutions to make their own decisions as to the content of their business curriculum (AACSB, 2021, 2013, p. 35-36).

Executive Coaching is Launched as an Industry. In 1995 Thomas Leonard started a non-profit called the International Coach Federation (ICF). His purpose was to support fellow coaches with an organization focused on coaching and research. He advanced the coaching profession by setting high standards, providing independent certification, and building a worldwide network of trained coaching professionals (ICF 2021; International Coaching Federation, 2020: International Coaching Federation Website). Additional professional and international coaching organizations that have strengthened the profession include The International Authority for Professional Coaching and Mentoring (IAPC&M), Center for Credentialing and Education (CCE), European Mentoring & Coaching Council (EMCC), International Association of Coaching (IAC), and the Association of Coaching (AoC).

The Growing Profession of Executive Coaching. Managers, leaders, human resource professionals, and talent-development professionals are using coaching skills and knowledge in their daily workplace interactions. A 2019 study described the range of coaching modalities and competences being used. The coaching continuum includes five categories: (1) Managers and leaders, (2) Human Resource and Talent development professionals, (3) individuals who work as *both* internal and external coach practitioners, (4) internal coach practitioners, and (5) external coach practitioners. Relevant to this study is that 99% of the responding coach practitioners completed coach-specific training. 85% of the training was accredited or approved by a professional coaching organization (e.g., ICF), yet only 10% of the training was university-based (International Coaching Federation, 2020).

Organizational Development. The intersection of managing, leading, and coaching supports the development of a coaching culture, which links to coaching and employee engagement, training and professional development for leaders and managers to learn coaching skills and leadership development for new managers and for millennials, as well as coaching to manage change. Increasingly, organizations plan to expand the use of coaching skills by their organizational leaders and managers. Respondents reported that coaching activities were used to develop leaders (55%), enhance performance management discussions (49%), and to develop talent (51%). The characteristics of high-performing organizations were correlated with organizations reporting a strong coaching culture, including high-performer retention (Filipkowski, Heverin, Ruth, 2019).

The Theoretical Foundation for the Coaching Profession. The ICF defines coaching as *partnering with clients in a thought-provoking and creative process that inspires them to maximize their personal and professional potential* (International Coaching Federation website, 2021), separate from mentoring, consulting, and therapy. The coach uses questioning to discover, clarify, and align with what the coachee wants to achieve. The coach encourages self-discovery, so that the coachee can generate solutions and strategies that move forward the coachee's agenda forward (Filipkowski, Heverin, Ruth, 2019, p. 5). The theoretical foundation for coaching builds on a rich history of self-help philosophers from the 1800s and early 1900s focused on positive thinking for mental gain. Collectively the emphasis was on self-reliance. Humanistic Psychology (1950s) led to understanding human drive and seeing the individual as good with a holistic view of raising human potential. Abraham Maslow was considered breakthrough in the field of human potential. Kurt Lewin's work on Group Dynamics helped individuals explore what happens inside groups. (Maslow, 1968; Wildflower, 2013). Mace and Mahler introduced managers to coaching skills by the 1970s and it began to take hold in the business world. Managerial coaching optimizes potential in a style of participative management and leadership, rather than an autocratic or directive style of management (DiGirolamo & Tkach, 2019).

Why Executive Coaching is Valued in Organizations. More organizations are hiring executive coaches to help onboard senior leaders, to remove barriers so that talented leaders are able to move up, and to optimize leaders' abilities and influence by removing limiting thoughts. More senior leaders are being trained to coach others within the organization. The ability to increase human potential through partnering with a coachee to change the way he or she thinks, is gaining recognition in organizational life. Yet, we allow the discipline of coaching to live outside of the traditional business curriculum? Why are coaching skills not being actively taught to our business students at a foundation level when they are undergraduates and at an advanced level within graduate school? Coaching directly aligns with and supports participative leadership and empowerment. It serves leaders and managers as they gain insights on their own thinking; it also serves leaders as they learn how to motivate and empower their followers.

INSIGHTS OF SENIOR LEADERS—REVIEW OF PART II (For full paper, see Mackenzie-Ruppel, 2022)

The insights gained from a primary stakeholder of business education was reported in PART II. The subjects for this study were ***senior leaders*** willing to share their experience and expectations of business education as pertains to executive and leadership coaching. The subjects included eleven senior leaders: 18% - President level, 37% - Executive Vice President/Senior V.P. level, 18% - CIO/CFO level, and 27% - Vice President/AVP level. Of the eleven senior leaders, 18.2% were female and 81.8% were male. A diversity of industries was representative with the Healthcare industry being the largest. The Part II paper provided details as to the research methods employed to gain the insights summarized here.

RESULTS

Fifteen (15) themes emerged from the research results and are presented under three umbrella concepts: (1) executive leadership coaching within organizations, (2) expectations of business graduates after being hired, and (3) expectations of business educators in preparing graduates to be workplace successful.

CONCEPT 1: Executive Leadership Coaching within Organizations. Seven themes revealed how executive coaching has become a valued tool for leader development. An intimate insight into the inner workings of leadership development was revealed.

Theme 1 - Leaders are exposed to executive coaching in different ways. Many of the leaders were exposed to coaching either as a recipient being selected as high-potential talent, or as an observer of others in the organization receiving coaching. For some, coaching was part of the day-to-day communications with other workplace colleagues or in the activity of managing interns. It was not always clear to the senior leaders as to how executive coaching *fit* into the workplace landscape; was it positive *or* negative that a person was given an executive coach?

Theme 2 - Shift in coaching from being perceived as remediation. Coaching has made a shift *away* from being perceived as a form of remediation. It is now viewed as a valued gift to those leaders projecting the most potential. Coaching is also used to help a valued leader to transition or onboard to *fit* into the new or current organizational culture. Coaching is also often used to upskill a newly promoted or hired senior leader. A senior leader may resist being given an executive coach because of the perception that coaching is a sign that you are *not* on track.

Theme 3 - Senior leaders align coaching with mentoring. When exploring how the senior leaders perceived the process and concept of coaching, its alignment with *mentoring* revealed a near full overlap. Mentoring expects that the coach has previously walked *similar steps* that the coachee is now needing to navigate, and therefore can provide guidance. In summary, the senior leaders align coaching most closely with mentoring with some seeing a broader overlap with consulting and even a little therapy.

Theme 4 - External vs. internal organizational coaches. Insight on the benefits and/or challenges of hiring either external or internal coaches to support leadership talent revealed the factors that CHROs must consider. "*Internal coaches will reiterate*" and reinforce the current culture, yet this is "*wrought with bias*." Hiring an external leadership coach will bring an "*external view*", yet an external coach is unable to help the coachee "*figure out how to fight for resources for your project … and how to handle tension.*"

Theme 5 - A coaching culture is desired, but it is unclear how it is defined. It was clear that working in a coaching culture is desirable, but there was no clarity as to how a coaching culture is defined. Senior leaders expressed that listening, communicating, collaborating, gaining insight, spending time with people to know them, and to become known, all contributed to developing and maintaining a coaching culture.

Theme 6 - Measuring the benefits of coaching. *How* to define the payback on a company's investment for an executive coach for a new, rising, transitioning, or sitting leader, is unclear. A direct link between the investment and the return is unrealistic. As one leader stated, "*coaching requires patience*."

Theme 7 - Coaching resources dedicated to support diversity, equity, and inclusion. An organization's commitment to diversity, equity, and inclusion is demonstrated by who is selected to receive the benefits of leadership coaching, which can make a significant difference in the career trajectory of the coachee.

CONCEPT 2: Expectations of Business Graduates After Being Hired. Three themes shift us from an insider view of coaching *within* the organization to an external view of what a graduate could expect as he or she transitions from student to employee.

Theme 1 - Who will get the job? Senior leaders are seeking the traits that will allow the candidate to integrate, thrive, and rise within the organization. Senior leaders desire new hires to be coachable, to demonstrate that they are team-oriented, hard-working, desirous to get involved, and willing to fit in.

Theme 2 - Expectations of young leaders as they develop and rise in the organization. Senior leaders are immediately assessing a new hire for current and future challenges and opportunities. An expected quality of rising leaders is "*self-awareness*." Rising leaders also need to "*be able to dissent*." They need to learn how to disagree, by supporting their positions, and "*saying it in a professional manner*." Other desired traits include the ability to *empathize, actively listen, and actively receive and integrate feedback*.

Theme 3 - Expectations of rising leaders as they begin to lead others. The natural progression for the employee is to lead others. This led to insights about the one-on-one interactions between supervisor and subordinate. *One* path revealed that the employee relationship is economic in nature. The employee is being hired for a job, and the explicit exchange of salary for results, should *not* be minimized. The *other* path revealed a coaching vibe that expected our rising leader to offer support and encouragement. It was suggested that "*a genuine interest in helping people do what they are trying to do*" be demonstrated. The use of powerful questioning to inspire subordinate development emerged in the data as valued.

CONCEPT 3 - Expectations of Educators in Preparing Graduates to be Workplace Successful. Five themes emerged that paint a picture of what business educators need to do to prepare students, *beyond* knowledge acquisition. They are (1) weakening student entitlement, (2) managing student expectations, (3) raising the need to gain real-world experience, and (4) acquiring skills that allow knowledge to be applied, and (5) *how* business education should *integrate* coaching skills and knowledge into the curriculum.

Theme 1 - Weakening student entitlement. Entitlement from an individual who has yet to earn it will alienate support from a senior leader. Some senior leaders believe that the entitlement is nurtured during the college journey. An inflated sense of entitlement for recognition for moderate work products, needs to be severely dampened. An entitled behavior can cause a young leader damage to his or her fledgling career.

Theme 2 - Managing student expectations. Dampening entitlement is tied to managing expectations. The senior leaders expect developing leaders to sit in many chairs and win many battles, before being elevated. Expecting rapid promotions may be a symptom of a college culture where praise is given to readily. Senior leaders expect faculty to set higher standards, forcing students to lift a heavier load before feeling success.

Theme 3 - Gaining real-world experience. Senior know that *experience*, as compared to education, can be a powerful teacher. Experience can be gained from professional work assignments as well as integrated into the student learning journey. It was suggested that forced failure be built into the classroom experience.

Theme 4 - Acquiring and using skills. Senior leaders expect their new hires to have learned how to *apply* gained knowledge. The ability to speak with ease and maturity, emerged as a sub-theme. The soft skills were consistently cited as necessary for workplace success (e.g., E.Q, empathy, working in teams)

Theme 5 – Integrating coaching skills, knowledge, and practice into business Education. Senior leaders recognized the value of coaching as *both* a skill and a disposition. These leaders were generous with their suggestions as to *how* coaching skills could be practiced during the college years. Some recommended a course or a concentration, but also recommended integration into the classroom.

DISCUSSION of Part II Research Results

In the past, executive coaching was perceived as remediation for a senior leader who may be getting too close to the cliff and needed to be rescued. This perception is fading as the most valued leaders are welcoming the partnership of an executive coach. It was made clear that senior leaders view coaching as an extension of mentoring, though the coaching profession suggests otherwise. The decision to staff internal coaches vs. hiring external coaches speak to culture (reinforcing it with an internal coach) vs. breakthrough thinking (seeing the broader pathway which an external coach may inspire). The concept of a coaching culture within the organization was raised but needs time to mature. The dilemma of measuring the coaching ROI benefits is a challenge that will not be easily resolved. And finally, the linkage of the coaching resource and its support of an organization's DEI initiatives found its way on to the radar of this research.

Senior leaders desire new hires to be *coachable*, team-spirted, hard-working, and willing to fit into the culture of the organization. After being hired, leaders will focus on behaviors that allow for meaningful and purposive interpersonal interactions. Self-awareness, empathy, self-understanding, the ability to disagree, and the willingness to be coached, are expected of the rising leader. And ultimately, if successful, the rising leader will lead others with confidence and skill.

This leads to our research goal of understanding what **business educators** can do to prepare graduates for success in the workplace. Senior leaders expect that faculty will lessen student entitlement, manage their expectations, and help them gain both skill and experience. The integration of coaching skills, knowledge, and disposition emerged as a needed component in business education.

INSIGHTS OF BUSINESS FACULTY – PART III

The objective of this research agenda is to explore the growing profession of executive coaching and to determine if its body of knowledge, skills, and theories should be integrated into business education. This series of papers has provided the landscape of the profession, its theoretical foundation, and the insights of those who benefit from the preparation of business students: employers and their senior leaders. Those who hire business graduates expect our students to be prepared to work in a coaching culture, be coachable, and to be able to coach others The insights gained from a primary stakeholder of business education, has led us to those who design and deliver this valuable education and professional preparation: Business faculty.

The intended outcome of this research will be generalizable knowledge related to whether Management and Leadership education needs to actively integrate into its program-level outcomes, the coaching knowledge, skills, and theories. The target audience of the final research results will be the business, management, and leadership scholars and academics who influence textbooks and curriculum within business schools. Business education must remain cutting edge. As we raise-up business education, we will better serve the world of business, both employers and future leaders.

Research Subjects for Part III

Eleven Business Faculty were interviewed from eleven different institutions of higher education. Of the eleven faculty: 45% were male, 55% were female; 45% held the academic rank as Full Professor and 55% held the rank of Associate Professor. The subject group reported a total of 223 years of college level teaching experience. The average teaching years of the eleven faculty members is over 20 years each with a range of 7 to 40 years. Even removing the outliers, the average teaching experience is 20 years.

All the faculty subjects worked within the management discipline, with many having an additional discipline, such as Accounting, Information Technology, Economics, Business Law, and Business Analytics. None of the faculty subjects are from the same college/university, so 11 institutions are represented. Of the 11 institutions, 9 of the 11 are in the state of New York, with 1 within a borough of the City of New York; 3 are located on Long Island, NY, and the remaining 5 being spread through the state of New York. The other 2 institution are located in NJ and KS.

Research Methods

In-depth interviews were used to capture the experience, insights, viewpoints, and opinions of the faculty subjects. The interviews followed the protocol for a semi-structured interview, which allowed the researcher to probe the subject's answers, while still maintaining the basic interview structure.

The interview schedule of questions inquired as to each faculty member's teaching focus and whether the faculty member's college had considered integrating coaching skills, knowledge, or practice into the curriculum. The professor's familiarity with the coaching process as compared to mentoring, consulting, and therapy was explored. What business educators can do to better prepare graduates to: contribute to a coaching culture, use coaching skills and knowledge in their development of others, and to be prepared to be coached as they rise through their careers, were explored. The professor's opinion was captured as to the

potential benefit and/or risk of integrating coaching skills, theories, and knowledge, into college-level business education, and what would be needed to accomplish the integration

The interviews were recorded and transcribed without any subject identifiers being captured (the IRB approved protocol was followed). The original recording was destroyed as soon as the transcript was created. The data was systematically collected and analyzed. Content analysis was used to collect the data points from the transcribed interviews. An iterative process was used to reduce the data and allow the themes and patterns to emerge. Grounded theory allowed patterns to form and the theories to be generated.

RESULTS

A total of 110 data points were systematically drawn from the eleven transcripts. These data points are verbatim sentences or sets of sentences. The data points were reduced in an iterative process to allow the themes to emerge. A total of eleven themes emerged that are being presented under two umbrella concepts, which are: (1) understanding coaching, and (2) the work of the faculty to deliver coaching skills and knowledge to students.

CONCEPT 1: Understanding Coaching

Many of the faculty had heard of coaching yet had a mixed level of acceptance and awareness. Some had no familiarity with coaching beyond what would align with coaches in the world of sport. As descriptors of what is *not* coaching were placed in front of the faculty, deeper insights and understanding emerged. The faculty started to align coaching with areas that made sense for them. A total of 25 datapoints drawn from the eleven leaders' interview transcripts, were reduced, allowing **five (5) themes** to emerge that reflect the faculty seeking to understand coaching.

Theme 1 – Faculty not aware of coaching as a professional business discipline. Though not a major theme within the results, it is still relevant to report that not everyone recognized that executive coaching exists in the world of business. For some, the lack of awareness did not prevent the faculty member from valuing its existence. Select data points that illustrate this theme include:

- ✓ *"Not that I'm aware of actual coaching."*
- ✓ *"Not formally familiar with a coaching model, but I think that this is what you do as a good management teacher or instructor. I think that's some of that comes from trying to work your hardest for the students to achieve great things."*

Theme 2 – Coaching brings to mind – sport. Though not a major theme it is relevant to faculty seeking and finding alignment with coaching as a business discipline. Initially some faculty had difficult separating coaching from sports. Some suggested that the word coaching has no meaning in business curriculum beyond "*the athletic arena*". One faculty shared his own mentor's reliance of influential athletic coaches' behaviors as examples to illuminate the difference between leading and managing. It was suggested that *"one way to think of coaching more directly resonates with the sporting context."* One faculty member used the example of athletic coaching to express that the coach does *not* need to have been the greatest athlete to be successful as a coach. Select data points that illustrate this theme include:

- ✓ *"Coaching is not a word that has any meaning in our business curriculum. People are familiar with coaching in the athletic arena."*
- ✓ *"My previous mentor was very passionate and would explain the difference between coaching and managing; he would always give us the examples of baseball managers and baseball coaches."*
- ✓ *"I would use a sports analogy in that a coach does not necessarily to have been the best baseball player but can still coach another to be great."*

Theme 3 – Coaching as compared to mentoring, consulting and therapy. The formal coaching profession makes clear that coaching is *neither* mentoring, therapy, nor consulting. When exploring how the business faculty perceived concept of coaching, there were *some* faculty that expressed "*surprise*" that mentorship was separate from coaching, suggesting that these concepts '*get muddled together*" and that "*they are all connected*." More often the faculty accepted the separation of these four concepts. This acceptance is in contrast to the senior leaders' perceptions; leaders more closely linked coaching and metoring. Mentoring suggests that the coach has previously walked the same or similar steps that the coachee is now needing to navigate, and therefore can provide the meaningful guidance or mentorship. The faculty more easily grasped coaching as a discipline and set of skills that can help a coachee inspire and shift thinking. Select data points that illustrate this theme:

- ✓ *"We will need help to understand the difference among coaching, mentoring, and consulting. Those all get muddled together; what's the difference?"*
- ✓ *"Mentoring means somebody else is going to throw them a rope and pull them through the situation."*
- ✓ *"was surprised when you said that they exist separately from each other because to me, they are all interconnected."*

Theme 4 – Communications as a proxy discipline and skill. The faculty naturally linked communications skills with coaching skills. As the path for success is envisioned, an individual with well-developed active listening and speaking skills is *also* envisioned. A connection emerged that improving communication skills is a prerequisite for coaching skills. It is suggested that a coach requires "*reflective or active listening skills*." Also, the student must minimize any predetermined idea and focus on "*asking the questions that get folks to thinking*." A prime focus was on listening skills blended with critical thinking skills that would inform powerful questioning. Select data points that illustrate this theme include:

- ✓ *"We talk a lot about the importance of communications as a path for being successful in any organization."*
- ✓ *"Before they can move forward with coaching others or being coached; active listening must be developed…. I think communication goes hand-in-hand with coaching."*
- ✓ *"Develop listening skills in a business communication class, they need reflective or active listening skills. Learn to not have a predetermined idea in mind, mentoring but coaching, It sounds like you need to start asking the questions that get folks to thinking. Help them practice those skills."*

Theme 5 – Curriculum needs to be reflective of what employers need. Faculty have a role in society as gatekeepers for their disciplines (e.g., humanities, sciences, or the professions). Unlike other disciplines, the professional world of business requires keeping a collective thumb on the pulse of what business leaders and employers need from their future employees. This theme reflects the bridge from the faculty seeking and gaining *understanding* of coaching *to* the next umbrella concept where the faculty work to *deliver* this knowledge and these skills to their students. Select data points that illustrate this theme:

- ✓ *"Making sure that those skills are what employers want so they don't have to train them when our graduates get there. We don't want employers to be disappointed."*
- ✓ *"We need to be aware of the changing landscape of business industries."*
- ✓ *"The type of curriculum that we're providing must be appropriate and driven in large part by what employers' need are; that relationship is key."*

CONCEPT 2: Work of the Faculty to Deliver Coaching Skills and Knowledge to Students.

The faculty moved from *understanding* to *action* as they recognized their role in delivering the knowledge and skills that the profession requires of graduates. A total of **85** datapoints drawn from the eleven faculty transcripts were reduced, allowing six (6) themes to emerge. These themes collectively tell a story. At first the faculty seeks where coaching skills and knowledge may already exist in the curriculum, next the faculty become more focused on where it *could* be inserted in the current curriculum, then they move to a clear affirmation that coaching needs to be *established* as a course, certificate, badge, or program. The story continues as they reflect on required faculty expertise in effectively delivering this new curriculum. The final theme is the realism as to the challenges that higher education faces in creating new curriculum.

Theme 1 – Faculty find where coaching may overlap the current curriculum. Faculty shared their realization that coaching already lives in parts of the curriculum. Their alignment of coaching with consulting, managing HR, and leading change, had faculty *seeing* the coaching knowledge and skills within those modules. Faculty named both programs and courses where they suggested that coaching already lives. This included such programs as the Master of Organizational Development and in such courses as entrepreneurship and organizational behavior. Faculty also reported that coaching skills are overlapping in the area of career development. Select data points that illustrate this theme include:

- ✓ *"We do not have a coaching class, but we have consulting, we have change management, we have HR, we have stuff that might straddle the border of this notion of coaching.*
- ✓ *"We have career services, that talk about resumes and interviewing skills and how to use online databases and the value of internships."*
- ✓ *"I feel that a fair amount of coaching happens both directly and indirectly; in our entrepreneurship courses as well, because the students create business plans for different, venture ideas, and you have to make sure that they are viable. And so, there's coaching, "*

Theme 2 – Faculty recommend places in the current curriculum where coaching may fit. As the faculty members' thinking expanded around the knowledge and skills of coaching, recommendations emerged as to where the natural fit may allow for *insertion* in the current curriculum. It was suggested that undergraduate students may be less ready to learn and demonstrate coaching skills. Rather, it was suggested that undergraduate students become exposed to "*being coached,*" and to learn "*what it feels like to be coached*" and experience "*the benefits of coaching*." Graduate students are better prepared to learn *how* to coach. The explicit courses that faculty recommended for the integration of coaching skills and knowledge included: Leadership, Organizational Behavior, and Human Resource Management. Career and professional development were also explicitly recommended as a natural fit where "*coaching at a more basic level would work …in terms of interview skills, resume writing, or communications."* It was also suggested that "*every one of our fields could include coaching. There could be coaching related to analytics, coaching related to project management, and working with teams."* This theme found faculty actively looking for the spots in the current curriculum that would welcome coaching skills and knowledge. Select data points that illustrate this theme include:

- ✓ *"In the organizational behavior class, you're talking about employee satisfaction. It's not framed in coaching language; but there would be overlap. We could easily change some of the language to talk about coaching in those classes."*
- ✓ *"So, my first thought would be that HR and OB professors could articulate the model that you talked about, you know, coaching, therapy, mentoring, and consulting"*
- ✓ *"It's the final week of the human resource class, but I left time to give them an overview on coaching, this is just like a value added. I'm giving them an overview on coaching just because it's not anything they've heard or know."*

Theme 3 – Affirmations made as to establishing a form of coaching as a course or program. As faculty consider where coaching currently show up in the curriculum, and then recommended where it could be inserted, the affirmations for establishing new courses, certificates, and programs emerged. Faculty recommended establishing new learning opportunities at the graduate level. One faculty member remarked that "*I don't really think our undergraduate students are ready. They have a hard enough time figuring themselves out*." Faculty recommended that coaching knowledge be delivered as a micro-credential, a stack-able certificate, or a separate course. It was suggested that coaching is "*too big a piece just to be sprinkled lightly into a class here or there. It's something that would need its own course*." It was also suggested that the coaching course be more fully integrated into the curriculum to ensure overlap to avoid having the coaching course be "*stranded by itself in a little corner and not allowing the students to get the added benefit*." Another professor recommended a "*dedicated track"* be developed within the MBA program. As the faculty members' thinking expanded, a theme of revenue opportunities emerged while citing the low cost and low risk that a coaching program presents. One professor stated that ***"y****ou could start this program with low risk, high reward."* Select data points that illustrate this theme include:

- ✓ *"If coaching could lead to a certificate, that might be something of interest. Allow different certificates to be added to our students resumes; it can be attractive to employers."*
- ✓ *"I think even a dedicated track whether it's an MBA with a concentration in whatever the knowledge and skills that would prepare one to go in that direction for the proper credentialing.*
- ✓ *"I definitely see the benefit of integrating coaching because it can be a revenue producer, in terms of return on investment. It is education in a field that is not yet totally flooded."*

Theme 4 – Why and how might coaching be offered to students. This theme emerged with the greatest strength. Faculty shared both *why* and *how* coaching should and could be offered to students. Leading with the *why* created the motivation for integrating coaching into the curriculum. It was suggested that "*being part of a learning organization … or a coaching culture is accepting that there's going to be failure*" therefore helping "*each other understand what we need to do to move forward,*" is important. Understanding that "*failing is an opportunity to learn*" is an insight that also emerged from the senior leader interviews. Helping students to "*change criticism into coaching*" elevated the learning objective and aligned with other faculty insights that "*education is not just about business management*," rather the students "*need to go out and be lifelong learners*." The value of coaching in helping individuals respect each other was articulated by one professor suggesting that *"DEI is part of coaching; it is to get people to listen to others and to respect others and to really get to know them better."* The *whys* continued with faculty suggesting that coaching "*would help people with the changing society"* and is "*leadership training at its finest."* It was suggested that coaching is *"like a personal Kaizen process … to continually improve."* As faculty recognized the value for students, the recommendations emerged that it be embedded in general education because it supports "*critical thinking"* and "*problem solving …where you're asking questions."* The benefit to the student is the development of a *"growth mindset. We need to teach them how to instill and appreciate that".* Select data points that illustrate this theme include:

- ✓ *"Failing is an opportunity to learn; that's a big part of learning. Helping students be aware that they're learning from their successes and the failures; then teaching students to be able to help each other."*
- ✓ *"We are trying to generate the leaders of the future, so I hope that I'm inspiring them."*
- ✓ *"Faculty must educate students about what coaching is; that is important. We can educate them on how to be receptive to being coached, as well as how to coach as they rise through their organization."*
- ✓ *"Keep focused on the growth mindset. We have to teach students not to perceive feedback as an attack; it can be supportive, it can be helpful, it can be something to build upon."*

Theme 5 – What do faculty need to offer a coaching course or program. As faculty became more committed to integrating a coaching curriculum, the question was raised as to what the faculty would need to feel confident and competent in teaching this curriculum. The data reflected the faculty desire to gain the needed knowledge, skills, and credentials. It was suggested that "*you need people who are cutting edge to teach.*" The faculty desire "*expertise*" in coaching before being expected to teach. It was suggested that "*current faculty develop that expertise … and to have expertise come from the outside, or a combination of both, which is typically what happens*." It was suggested that faculty be supported to gain licensure or the coaching credentials that are recognized by the profession. The faculty *"want to fully understand what coaching is."* Select data points that illustrate this theme include:

- ✓ *"They should get me some licensure … before I can teach coaching."*
- ✓ *"So as a coach-teacher, I need certifications; I need all of the curriculum that is involved in the coaching body of knowledge, I would want to fully learn and fully master that."*

- ✓ *"I think it would it makes sense to have faculty educated on coaching."*

Theme 6 – Barriers to establishing a coaching curriculum. The final theme expresses the barriers to curriculum creation. This is a valuable theme as it may help us understand why this growing profession is gaining its education *outside* of the colleges and universities. Though the faculty may be willing, the process to develop and offer new curriculum appears to be a roadblock. Barriers included the lack of qualified faculty; "*I don't know how many people have PhDs in this field".* Another barrier is motivating faculty to change suggesting that "*it is a fairly heavy lift and hard to get faculty to agree to change curriculum*" and that "*it's hard to reach consensus*." Another barrier is convincing administration that the market demands the change, with one faculty member remarking that his "*school wouldn't start coaching unless they heard executives talking about coaching*." Another professor stated that *"…must be able to demonstrate that there is demand in the market for this sort of thing for it to be supported."* Another professor shared that the barriers are less economic and more political, stating that "*I think the politics of the college is too much; why are you knocking your head against the wall with colleges?"* And the final barrier that was shared was the low motivation of senior faculty to learn a new field; "*I'm past that at this point in my career"* and "*It's kind of late for me."* One professor stated that *"There's a lot of teachers out there that have apathy. And it's a lot easier to just show the PowerPoints that come with the textbook.* Select data points that illustrate this theme include:

- ✓ *"It's just so hard to make new curriculum happen; realities of developing curriculum and getting it approved."*
- ✓ *"Getting the right staffing resources, is the first thing to ensure, which is a real challenge…"*
- ✓ *"A year ago, we had this little contest to develop new programs for the college. And I submitted a minor in organizational coaching. But it didn't win, although sustainability won because it got the buzz."*
- ✓ *Building this new coaching program will take a lot of energy and a lot of work."*

CONCLUSION OF PART III PAPER

A transformational mindset is needed to ensure that business educators are aware of what is needed and desired in the marketplace so that there is ***no gap*** in the preparation for these rising professionals. This must be an ***active***, rather than passive process. As the executive coaching profession has grown, there appears to be only minimal evidence that business education has been updating its curriculum and learning outcomes to incorporate what is expected. The insights from the Part III research reveals that faculty support the integration of the coaching skills and knowledge into Business education. They also shared the barriers and roadblocks that will need to be overcome.

NEXT STEPS FOR PART IV and PART V OF THIS MULTI-PHASE STUDY

The exploration will continue with interviews of our final stakeholder: MBA graduates who have now moved fully into their professional roles. The **Part IV** research will invite in-depth interviews from eleven MBA graduates. Their reflection on what they experienced and how their MBA education *did* or *did not* prepare them for the workplace, will provide the needed triangulation for this research agenda. **Part V** will draw together the stakeholder insights with the intention of developing explicit recommendations for improving business education. Part V will also include an environmental scan of how the skills and knowledge of the executive coaching profession is currently showing up within business education. The goal of this exploration is to not only answer the question, *Does Management Education need a facelift?* But to make meaningful recommendations that will serve both our students and stakeholders.

WORKS CITED

AACSB. (2013/2018). "2013 Eligibility Procedures and Accreditation Standards for Business Accreditation." URL: https://www.aacsb.edu/-/media/aacsb/docs/accreditation/business/standards-and-tables/2018-business-standards.ashx?la=en&hash=B9AF18F3FA0DF19B352B605CBCE17959E32445D9

AACSB. (2018). "AACSB Industry Brief: Lifelong Learning and Talent Management." URL: https://www.aacsb.edu/-/media/aacsb/publications/research-reports/lifelong_learning_paper_final.ashx?la=en&hash=BA7D25D6C2BCBEA905C2ACFB5F3F80BEDE64A6C7&_cldee=bW1hY2tlbnppZS1ydXBwZWw1QG1vbGxveS5lZHU%3d&recipientid=lead-b06334b166f9eb1194ef000d3a1394de-df09f93b60c6476c9fa2a98c7599a0e4&esid=60d7ecc6-2da4-4b49-ad1b-7fcb3283cca7

AACSB. (2020). "Guiding Principles and Standards for Business Education." URL: https://www.aacsb.edu/-/media/aacsb/docs/accreditation/business/standards-and-tables/proposed%202020%20aacsb%20business%20accreditation%20standards%20-%20final%20draft%20-%20april%206%202020.ashx?la=en&hash=B40646D6F0057FBAF289B3B04888A33BB2741A3D

AACSB. (2021). "How AACSB is Transforming Business School Education." URL: https://www.aacsb.edu/accreditation/resources/white-papers

DiGirolamo, J.A., & J.T. Tkach. (2019). An exploration of managers and leaders using coaching skills. *Consulting Psychology Journal: Practice and Research*, 71(3), 195-218.

DiGirolamo, Joel A., and J. Thomas Tkach. (2020). "An Exploration of the Coach Approach to Managing and Leading: A white paper for managers, leaders and professional coaches". International Coaching Federation (ICF).

Filipkowski, Jenna, Abby Heverin, and Mark Ruth, (2019). "Building Strong Coaching Cultures for the Future. Final Report". Human Capital Institute.

Harcleroad, Fred G. (1980). "Accreditation: History, Process, and Problems". American Association for Higher Education, Washington, D.C.: ERIC, Clearinghouse on Higher Education, Washington, D.C:

Harcleroad, Fred G., and Judith S. Eaton. (2005). "The hidden hand: External Constituencies and their Impact," Chapter 9 in Philip G. Altbach, Robert Oliver Berdahl, and Patricia J. Gumport (editors). *American Higher Education in the Twenty-First Century: Social, Political, and Economic Challenges*. JHU Press. ISBN 0-8018-8035-1. Pgs. 262-264.

Hegji, Alexandra. (2018). "The Higher Education Act. A Primer". Congressional Research Service 7-5700 www.crs.gov R43351

International Coaching Federation Website. (2021). URL: *CoachingFederation.org*

International Coaching Federation (ICF). (2020). "2020 ICF Global Coaching Study. Executive Summary". URL: https://coachingfederation.org/research/global-coaching-study

International Coaching Federation (ICF). (2021). "History of ICF". URL: https://coachingfederation.org/history

Jamison, Jamie L. (2018). "The importance of front-line supervisor coaching ability and performance to employment decisions and coaching culture success". Dissertation. ProQuest number: 10829429

Mackenzie-Ruppel, Maureen L. (2020). "Higher Education Underperformance, PreCovid19! *Proceedings of the 47th Annual Meeting of the Northeast Business & Economics Association*. November 12-14, 2020. Hosted by Peter T. Paul College of Business and Economics, University of New Hampshire. Print: ISSN 1936-203x Online: ISSN 1936-2048, pages 103-107.

Mackenzie-Ruppel, Maureen L. (2021)." Does Management Education Need a Facelift? The Intersection of Managing, Leading, and Coaching. Part I?" *Proceedings of the 48th Annual Meeting of the Northeast Business & Economics Association*. November 4-6, 2021. Hosted by Stockton University, Stockton, NJ. Print: ISSN 1936-203x Online: ISSN 1936-2048, pages 147-153.

Mackenzie-Ruppel, Maureen L. (2022, in press)." Does Management Education Need a Facelift? The Intersection of Managing, Leading, and Coaching. Part II?" *Proceedings of the 49th Annual Meeting of the Northeast Business & Economics Association*. November 10-12, 2022. Hosted by University of New Hampshire. Print: ISSN 1936-203x Online: ISSN 1936-2048.

Maslow, Abraham. (1968). *Toward a Psychology of Being*. Wiley & Sons, Inc. ISBN: 044-20-3805-4.

Program Registration Guidance Documents. (2020, February 24). Department expectation: Curriculum. *New York State Education Department.* URL: http://www.nysed.gov/college-university-evaluation/department-expectations-curriculum

The Trends Report. (2019). *Chronicle of Higher Education*. URL: https://www.chronicle.com/article/the-2019-trends-report/

Wildflower, Leni. (2018). *The Hidden History of Coaching*. McGraw Hill Open University Press.

Teaching Business Statistics in a 2-Hour Lecture

ABSTRACT

We present a practice of teaching statistics to business students in a two-hour lecture, which covers topics of statistical inference, hypothesis testing, regression analysis, determining minimum sample size, and comparing multiple populations. Confidence interval is used as the key thread running through statistical analyses. Mathematical details and formulas unnecessary for business students are avoided. We focus on the essential concepts of statistics and on solving practical statistical issues in business. This lecture of statistics has been given to students multiple times in the past five years. The results of outcome assessment have showed that this two-hour lecture enables students to deal with major statistical applications in business.

We present our pedagogical practice of teaching business statistics in two hours to undergraduates and graduates.

I. INTRODUCTION

Business Statistics is a required course for business students, focusing on applications of statistics. It requires student's understanding of the essentials of statistics and skills of doing statistical analysis in varieties of business scenarios. Statistics is used in other business courses as marketing, finance, management, and business analytics. However, business students are typically weak in statistics. Trying to provide a re-education highlighting the essentials of statistics is the initiative of developing this 2-hour lecture.

Statistics is a challenging course for business students. It is hard because of the mathematics involved and students' weakness in math. They were usually lost in the details of concepts, probability distributions, and mathematical formulas in a traditional statistics course, and missed the essentials of statistics. At the beginning of this lecture, almost no student could tell "what statistics is about" and "why probability theory is needed in statistics", and no one could tell what "supporting rate 43.5%±3%" exactly means as the result of a poll, even though they took a statistics course before.

The facts are: (i) A lot of mathematical details and delicacies are not concerned in an application. (ii) Most of calculations in statistics can now be done in computer (Excel, for example). Therefore, a business student can use statistics in applications without knowing the mathematics about t-distribution, χ^2-distribution, and F-distribution, without knowing formulas for confidence interval and for minimum sample size, and without knowing method of least square and ANOVA.

This lecture does not need a textbook. Lecture slides are placed online. The lecture does not require students to have taken a statistics course before. High school algebra is sufficient to understand the contents of the lecture.

The goals of the lecture are set up as:
(1) Students grasp the essential idea of statistics.
(2) Students differentiate statistics from other subjects such as probability, database, optimization approaches, data processing, and data mining.
(3) Students understand the result of statistics analysis.
(4) Students are able to do statistics work in business applications, such as estimate population from a sample, doing hypothesis testing, using regression to tell the relationship between variables in the sample and in the population; finding minimum sample size for a particular case; comparing multiple populations as ANOVA does.

We use confidence interval, CI, as a thread throughout various statistical analyses. Following topics are covered in the lecture:
(1) Sample vs. population,
(2) Confidence interval, CI, and its calculation with the help of computer,
(3) Statistical analysis on population mean and population proportion,
(4) Hypothesis testing by using CI,
(5) Searching for minimum sample size in Excel,
(6) Solving ANOVA problems with CI,
(7) Regression equation and relation between variables, R^2, in a sample,
(8) Relation between variables in population, CI and *p-value*.

Most of students in class, after taking this lecture and doing assignments/quizzes, passed the exam containing application problems from business statistics textbooks in the fields of statistical inferences for population mean and population proportion, hypothesis testing, determining minimum sample size, statistical inferences for two or more populations, analysis of variance, and simple/multiple regression.

II. OUTLINES OF THE LECTURE

The essential idea of statistics is to investigate a population from a sample of it. CI, confidence interval, is the common thread running through all statistical analyses. Here are the outlines of the lecture slides.

Slides #1 to #4.

"Sample" is the key word of statistics

We do not need statistics if we want to know how many female students currently in the classroom, - We simple "count".

We have to use statistics if we want to know how many people in US smoke, because we are not able to "count" one by one. We need a "sample".

Slide #5.

Definition of statistics:
Statistics is a subject of investigating what the population is like from a sample of it.

Slide #6.

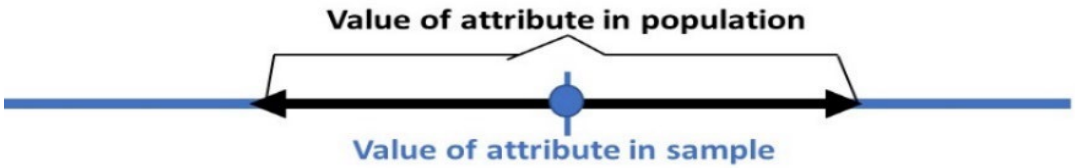
Attributes of a population that are most concerned in business applications:
Mean of population,
Proportion in population,
Relationship between variables in population.

Slides #7 to #8.

From a sample to its population:

The value of an attribute of a sample can be calculated by using the data in the sample.

The value of the attribute of population is very much likely not the same as the sample. Although it is unknown, but it should be near the value of the sample attribute.



The range where attribute value of population may stays is called *confidence interval*, CI.

Accurately saying, we believe with (1-α)% confidence that the population attribute value falls in the CI, where (1-α)% is called *confidence level* (by default it is 95%). In other words, probability that the population attribute value is in CI is (1-α)%, and probability it is not in CI is α%.

The higher the confidence level (1-α)%, the larger the CI. If (1-α)%=1, the CI is infinitely large, ∞.

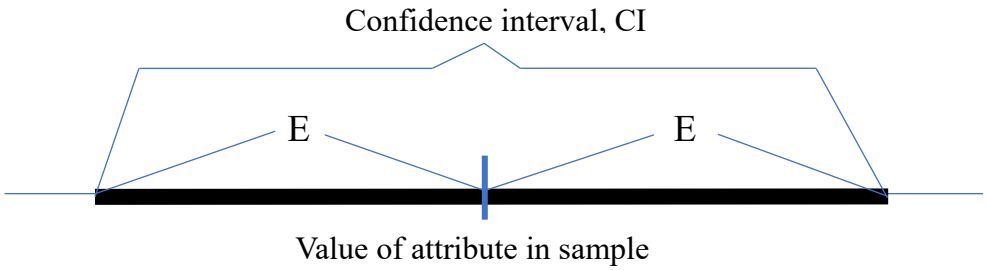
All statistical analyses can be addressed in terms of CI, as to be addressed below.

Slides #9 to #11.

Calculations of CI:

Probability theory is used to calculate CI, given the attribute value of a sample, standard deviation of the data in the sample, and sample size.

Let E = half of CI. That is, CI = 2*E.



An Excel function is used in calculation, which is E ***=confidence.t(α,σ,n),*** where α=allowable error rate which is the chance that the population mean is NOT in CI, σ=standard deviation of data in sample, n=sample size.

Slides #12 – #22.

Examples showing where to get CI.

(1) Population attribute: Mean.

Let M be sample mean, μ be population mean.

Population mean μ is somewhere in the CI = {M-E, M+E} with (1- α)% of confidence.

(2) Population attribute: Proportion.

Let P be the sample proportion, π be the population proportion.

Population proportion π is somewhere in the CI = {P-E, P+E} with (1- α)% of confidence.

(3) Population attribute: Relation between Y and Xs.

In a regression equation $Y=b_0+b_1X$, the coefficient of X tells the relation between Y and X. Y is negatively related to X if $b_1<0$. Y is positively related to X if $b_1>0$. Y is not related to X if $b_1=0$.

Regression equation $Y=b_0+b_1X_1+ b_2X_2$ is derived from a sample (by using the least-square method by computer).

Let $Y=\beta_0+\beta_1X_1+ \beta_2X_2$ be the regression equation in the population from which the sample comes from.

β_0 is somewhere in (CI of b_0),

β_1 is somewhere in (CI of b_1),

β_2 is somewhere in (CI of b_2).

where CI, defined as {lower bound, upper bound}, is provided in the results of each software of regression analysis.

Slides #23 - #30.

Examples showing how to use CI for statistical purposes.

(1) Statistical analysis:

CI indicates, with (1- α) confidence, where population mean is, or where population proportion is, or what the relation between variables in population is.

(2) Hypothesis testing:

To answer the question: In what case can we reject the null hypothesis H_0=(attribute=x in population)?

- Reject H_0 if x is outside the CI.

(3) Finding minimum sample size:

Search the minimum sample size in Excel by holding values of α and σ and keeping changing value of n in the function E=confidence.t(α,σ,n), until the function returns the designated E value.

(4) Compare two or more populations in terms of mean and proportion (as by ANOVA):

If two CIs do not overlap, then the two populations are significantly different in terms of a predetermined attribute.

(5) Relation between variables:

(i) Regression equation indicates the relation in a sample.

(ii) CI for a variable X_i , i=1, 2, …, indicates the relation in the population.

(iii) Hypothesis testing: We do not reject the null hypothesis that Y is not related to X_i in population only if CI of X_i contains 0, i=1, 2, ...

Slides #31-#33.

Auxiliary indicators in regression analysis:

CI is sufficient in most cases of statistics analysis. The other indicators provided in regression software can be viewed supplementary or auxiliary. For example,

p-value of X_i, which is the error probability if we reject the hypothesis H_0=(Y is not related to X_i). p-value<α if and only if CI with (1- α) confidence does not contain 0.

Significance F, which is the error probability if we reject the hypothesis H_0=(Y is not related to any of Xs). Significance F<α if CI for all β_i with (1- α) confidence does not contain 0, i=1, 2, ...

t-stat, so that |t-stat|>1 if p-value<0.05.

III. AN EXAMPLE OF SLIDES:

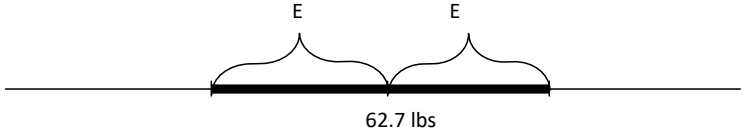
Here is a sample of lecture slides, showing the details of calculating confidence interval and how to use CI to do statistical inference by using an example of "South Jersey Boys".

Slides 12 – 15. **Investigate Population Mean**

Example: We want to know the average weight of boys between 5 and 7 in south Jersey.

Suppose that a random sample of 100 boys is picked from the population, and the average weight of those 100 boys is 62.7 lbs with standard deviation σ=7.48.

So, the population mean should be somewhere around 62.7 lbs.



In this example, sample mean is 62.7 lbs, σ=7.48, and n=100.

If α=0.05, then

E = CONFIDENCE.T(α,σ,n) =CONFIDENCE.T(0.05,7.48,100) = 1.4842.

CI for population mean: (62.7-1.4842, 62.7+1.4842) = (61.2158, 64.1842).

Interpretations of the range of CI (61.2158 lbs, 64.1842 lbs):
It is the 95% confidence interval that contains the population mean;
The population mean is in this range with probability 0.95;
The chance that population mean is not in the range is 5%.

Can we say that the average weight of boys between 5 and 7 is over 61 lbs?

Yes. Since CI (61.2158 lbs, 64.1842 lbs) is entirely over 61, we can say with 95% of confidence that the population mean of average weight of the boys is over 61 lbs.

What if we set allowed error rate α=0.01 (confidence level is 0.99)?

E for 99% confidence interval is:

E = CONFIDENCE.T(0.01,7.28,100) = 1.9646.

99% CI is:

(62.7-1.9646, 62.7+1.9646) = (60.7354, 64.6646).

Can we say that the average weight of the boys is over 61 lbs?
No, since 61 is within CI.

IV. OUTCOME ASSESSMENT

Students' feedbacks, such as questions in class, submitted assignments, quizzes, and exam, have told that they understand and catch the essential idea of statistics and technical details as we covered.

The effectiveness of this intensive statistics lecture was assessed by exam, shown as below. The exam covers statistical inference, hypothesis testing, determining minimum sample size, comparing multiple populations, and regression analysis.

Problems in Exam:

1. (20 pts) The average time required to assemble a unit in an assembly process has been 10 minutes. The most recent periodic sample of 22 randomly selected units reveals the following assembly times (in minutes):
10.8, 10.4, 10.6, 9.8, 9.7, 9.9, 10.4, 9.8, 9.6, 10.7, 10.3, 9.6, 9.9, 11.2, 10.6, 9.8, 10.9, 10.4, 10.5, 9.7, 9.9, 10.8 .
Does the sample significantly suggest that the current average assembly time deviates from the required 10.00 minutes (assuming 95% of confidence level)? If yes, what is the chance of error of the statistical suggestion?
(*Show your arguments, calculation steps, and conclusions, as well as Excel formulas (functions) and their parameters you use.)*

2. (20 pts) The advertising director for a fast-food chain would like to estimate the proportion of high school students who are familiar with a particular commercial broadcast on television in the last month. A random sample of 400 high school students indicates that 140 are familiar with the commercial.
(2.a) Set up a 98% confidence interval for the population proportion of high school students who are familiar with the commercial.
(2.b)The goal of the advertisement was to have at least 30% of high school students know the product. Can we conclude that the goal has been reached according to your result in (3.a)? Explain briefly. If we can conclude it, what is the chance of error of conclusion?
(*Show your calculation steps, Excel functions and their parameters you use.)*

3. (10 pts) The defective rate of a component from a production line is estimated at 3%. If we want to take a sample to statistically determine the actual defective rate of the components from the production line, then what would be the minimum sample size to make the size of 95% confidence interval less than 2% (i.e., E<1%)? Describe briefly how you derived your answer, what Excel function(s) you used and how you used the function to figure out the solution.

4. (15 pts) Use datafile: ***Home Market Value***

Develop a multiple regression model in which "market value" is taken as Y, "house age" as X_1 and "square feet" as X_2.

(4.1) Run Excel and find the regression equation.

(4.2) What is the value of R^2? What does the R^2 value mean?

(4.3) What is the CI of the coefficient of X_1 and X_2? What do they mean?

(4.4) Is Y related to X_1 significantly? Why or why not?

The average score of the exam was 53.78 out of 65, or 82.7%. The test results showed that students were capable of doing statistical analysis and regression analysis in business application problems. The preset goals of this lecture are largely reached, though there is room of improvement. "Statistics turns out not as scary as I used to feel," a student commented.

V. SUMMARY WORDS

Statistics is a branch of mathematics. It is different from other math subjects. Calculus, algebra, geometry, trigonometry, probability theory, and operations research are "deduction", in which general theorems are deduced to cover, explain, prophesy individual and particular occurrences. Statistics is "induction", which summarizes individual and particular occurrences and comes up with general assertions. Statistics goes bottoms-up, from specific to tell what the general is like; while other subjects go top-down, from the general theories to tell what specifics are like.

Statistics is complicated mathematically. But it is conceptually simple. It is to investigate a population from a sample of it. There will be chances of making errors when extending the result of a sample to the population. Probability theory is therefore used to tell the chances of errors.

Business people use statistics as a tool in their career. They do not need to understand every formula and every mathematical detail of statistics in career. And they have computers to help do calculations if needed.

As a mathematical subject, statistics is rigorous in dealing with various scenarios. In many applications of statistics, there contains inaccuracies and uncertainties. Business users may not be concerned with mathematics accuracies and subtleties, which usually turn around to distract the attentions from understanding the essential of statistics.

This two-hour lecture on statistics is a pedagogy renovation of business statistics. The feedbacks of students have showed that students have learned from this lecture, and business statistics can be effectively taught in this way different from the traditional textbooks.

Reference:

[1] Canavos, George C. and Don M. Miller, Modern Business Statistics, 1995, DuxBury Press, ITP

[2] Evans, James R., Business Analytics, 3rd edition, 2020, Pearson
[3] Levine, D., M. Berenson, D. Stephan. Statistics for Managers – Using Microsoft Excel, 8th edition, 2016, Prentice-Hall, Inc. Upper Saddle River, New Jersey

United States Postal Services

Background

The United States Postal Service, also known as USPS, is an independent agency within the US federal government. Their overall task was providing efficient postal services throughout the US, which can be traced back to 1775 during the Second Continental Congress. The Post Office Department was created after the Postal Service Act was passed in 1792. They employ nearly one-third of the federal civilian workforce while providing delivery services to over 130 million households and businesses. The Congress and public expect the agency to fulfill its primary objective of providing universal postal services at reasonable rates. With their affordable rates, they dominated, but of course, the postal service faced competition from private delivery companies and digital platforms while offering electronic communication alternatives. In recent years, the agency has reported long-standing performance and management challenges that must be addressed to ensure the postal service remains sustainable and competitive in the 21st century. The challenges facing the USPS for years have perplexed leaders and stakeholders of the industry. With new tech companies, the agency loses billions of dollars while entering the shipping and logistic market. The majority have called for its privatization, which would raise several political questions and require Congress to allow private entities to manage the second-largest employer (USPS). Therefore, in this case study, we will highlight some of the problems and challenges that the USPS is currently facing with suggestions of possible solutions that can help improve the service delivery in the government agency through information technology.

The internet has been a significant factor in how businesses operate. Roughly 90 percent of the American population has access to the internet, especially those individuals who are in cities and densely populated areas. Mobile access to the internet is equally on the rise, with public areas providing access to Wi-Fi. Hence, internet usage has risen across all age groups, not

just adults but the younger generations as well. As previously mentioned, the internet has transformed how businesses operate, allowing technological innovations to change rapidly in the e-commerce sector, especially within the United States Postal Service. With the increased competition in the e-commerce field, it has since provided incentives for market players (Finger & Bukovc 2014). For decades, the American postal service has provided American homes with mail and parcel delivery services. However, the operations of the government agency are yet to be in line with the current technological advancements fully.

Current Problems and Challenges

The real competitive advantage of the Postal Service is that it is the only organization embedded in every neighborhood in the United States. Therefore, its impact is felt across America when the agency faces trouble. The postal services' success depends on the agency's ability to control operating costs, digitalize service delivery, and ensure service integrity. In the recent past, the agency has recognized the need to become aggressive in its cost management and improve the internal operating controls to limit unnecessary costs. This is where the USPS faces growing challenges in a competitive environment as it strives to fulfill its mission. At the beginning of the 21st century, the agency was nearing its end with the emergence of digital platforms and the internet, which rendered traditional communication services unvaluable. Private delivery companies and electronic communication alternatives such as the internet have challenged USPS. First-class mail volume has declined substantially as more Americans diverted to electronic communication alternatives in the recent past.

Postal service enjoyed 100 percent of the Express Mail market for years, but the shares have been dropping yearly. The service has also been experiencing a similar decline in the parcel delivery market. Most USPS competitors have the freedom to select markets they can serve and

set prices based on the demand and cost of the service they provide. On the contrary, the Postal Service, by statute, is mandated to provide universal benefits to all communities. The service provided is also constrained by law while allocating costs in the different services and setting prices based on the market conditions. The noncompetitive postage rates have led to a reduced market share. In the international market, the USPS has the flexibility of setting its rates; however, the volume growth experienced has been considerably low. Customers often prefer the more reliable service providers, private companies with integrated systems.

USPS's financial outlook has deteriorated, with the agency experiencing significant losses in most of its fiscal quarters. The fiscal year 2019 deficit was reported to be roughly $5 billion, which may sound nice, but it means the agency had failed to meet most of its targets. The USPS operates on a cost-plus structure; hence the service adds up all the costs incurred and later adds markup to calculate the profit. The system is not effective enough as the customers pay for the inefficiencies. The structure also means that at any given point, USPS is unlikely to rationalize or bring down their costs but instead tries to push the price down to the consumers. The structure favors USPS in areas with no competition, such as postage. Still, on other services such as package delivery, consumers will go for alternative service providers such as UPS and FedEx, making it difficult for the USPS to enforce the cost-plus structure (Hillebrand et al., 2016). In the market of parcel handling, USPS are price takers and not price makers. The post-service has recently been inefficient with a very high fixed cost, which is why the service has been losing money.

Cause of the Problems

The problems that USPS has been facing stem from various reasons. First has been the agency's management, which has often been marred with incompetency. USPS has been

hampered by challenges in labor-management relations which have often resulted from the autocratic leadership seen in top management, the adversarial attitude displayed by their employees and direction, and the use of an inefficient performance management system. The management problems can be seen in how mail carriers and other employees get treated. As the second-largest employer in the United States, USPS is not providing its workers with the best working experience, which can be the reason behind the decline in service delivery. This is a huge issue that keeps getting ignored. For example, the inefficiency in time clocks for punching in and out has been inaccurate for many years. It is 10-15 minutes ahead, leading this to be another issue factor between management, staff, and HR. For some reason, the agency has found it challenging to retain most of its top employees and goes for less skilled individuals unaware of this issue. The delays witnessed in mail delivery result from a limited workforce and incompetency on the part of the workforce.

Secondly, USPS has always been slow to adapt to the changing times, especially the use of technology, as previously mentioned. Like many traditional businesses, the USPS saw the internet revolution as a fluke and one that would not last. Therefore, not prioritizing the use of technology in the early days is one of the reasons the agency lags in terms of efficiency and service delivery. Most of the services offered at USPS are manual and often cumbersome, leading to mistakes and delays in fulfilling orders. In comparison, their rivals such as FedEx, Amazon, and UPShave improved their services by adapting cutting-edge technologies enabling visibility and efficiency in service delivery. With these rivals' improvements, Americans can see their mail and packages delivered with the click of a simple photo, unlike USPS, which simply provides text-based alerts upon delivery.

Political interests and political interference have also been a challenge affecting the USPS operations. The USPS is an essential tool in the United States' political process because the US has a mail-in voting system. Voters can have the ballot papers mailed to them, and they vote from the comfort of their homes, work, or class and send them back. The postmaster is a presidential appointee, who adds to the neutrality controversy in the whole process. It should, however, be noted that the agency does not receive taxpayers' funding. The 2020 elections saw how political interference has been affecting the agency's operations, with one side of the political divide opposing additional funding being channeled to the agency. However, the ties between the Postal Service and politics can be traced back to 1792, when the United States founding fathers George Washington, James Madison, and Benjamin Rush utilized the postal network to inform the electorate. The connection between the political figures and postal services then was positive, and it was used for the nation's benefit. We all know the effect of the Postal department on politics, but this is where the problem arose, when the politicians began using the USPS for their benefit.

A good example is President Andrew Jackson (1767-1845), who created the spoil system and made the postmaster general a powerful Cabinet officer while installing his cronies to that position. The spoil system meant that any party that won the White House would reward its supporters with tens of thousands of jobs at USPS. However, the current crisis may be dated back to the 1980s when the USPS management and Congress declined not to shift to email. As a result, they had an opportunity to provide US residents with digital addresses similar to the traditional physical street addresses. Things worsened when the Postal Accountability and Enhancement Act of 2006 were passed, thus restricting the ability of the postal services to offer

new services or make adjustments to their prices (Kamarck, 2009). Those moves have since made the postal service the worst performing government agency.

Solutions to the Problems

Over the last two decades, technological innovations have advanced, and more improvements are still in process, which has enabled the creation of opportunities in almost all sectors and, at the same time, caused disruption in some industries. The postal sector is no exception, and to move with the current times, the agency needs to adjust and adopt some of the new advances. The USPS collects a considerable amount of data daily; however, the main challenge has always been putting the data into use and enhancing it to improve customer experience and counter competition from the emerging digital-based companies such as FedEx UPS and Amazon. The first solution to the problems that the USPS is facing is committing to a digital strategy capable of transforming mail and further monetizing different assets of the postal service. In the recent past, USPS has employed digital technology in optimizing its logistic network and delivery services which are vital for cost-cutting (Isa & Kiumarsi 2019). However, there are more attractive adjacencies that the agency can use to bring it back to profitability. In the world of digitalization, data is critical. Using information technology, the agency can build a data-rich environment for its customers, which will help accelerate the digitalization of physical products and further speed up a digital transformation that will ensure operations are optimized.

USPS can benefit immensely by having a robust data and analytic platform that will help aggregate and track data and analytics across the postal service supply chain and further help monetize the data for private use. The data analytics will help the agency offer its customer information, enabling greater visibility during delivery, which is a challenging aspect of the supply chain (Trimble, 2016). The robust data, tracking, and analytics will enable the

commercial customers to coordinate personalized mail to the consumers and to create further omnichannel marketing and communication tools (Fuglsang & Møller 2014). The USPS will also be able to use real-time data to provide business insights and further enhance customer value. Therefore, the agency can improve its payment, returns, and visibility of services. Postal infrastructure heavily relies on big data across the supply chain. The agency transports millions of emails and parcels using its various networks daily. Data gets generated through tracking the location of mail, the speed at which it gets to its destination, and the network helps identify problems. The web is often vast; hence the collection and utilization of such information can only be harnessed through data analytics. Data and analytics are core to the operation of companies and organizations in the information technology age as it helps improve efficiency and quality service delivery (Buko et al., 2022). The USPS is set to be more competitive in the process as it will enhance the quality of service.

Secondly, USPS needs to improve and explore ways to use mobile applications to help inform deliveries. The mobile application can be engineered in a way that it tracks and photographs letters, packages, and catalogs which are then shared with the users through the periodic update. An informed delivery system will effectively increase the value of direct mail as it doubles the number of impressions a user sees. The customer is presented with a digital and physical appearance. Therefore, marketers can use double impressions to target their address market. Customers are engaged daily and often in real-time, providing marketers more chances to reach consumers. A combination of physical and digital strategies is vital for enhancing growth in the sector.

Improved visibility in processing and delivery issues is critical for shipping and international deliveries. Think about it: competition such as Amazon can update the customer

with delivery confirmations via text or email and by photo for accuracy. Efficiency in tracking and managing freight using Radio Frequency Identification (RFID) eliminates the need to use barcodes in the scanning process. Data analytics also helps in the area of optimization of routes. Therefore, the agency can perform dynamic routing using real-time analytics that enable better allocation of delivery resources and loading. Areas of Artificial intelligence (AI) and Machine Learning (ML) are also areas that the agency can explore to help with predictive analytics for efficient allocation of human and financial resources (Marr, 2019). Continued development and implementation of tools that rely on analytics would eventually make the USPS more competitive. Advanced computer systems will help in offering streamlined services. Currently, USPS uses the Edge Computing Infrastructure Program (ECIP), a distributed AI system capable of processing 20 terabytes of images daily from 1,000 mail processing units. Building an enterprise-wide AI program for the USPS will be a motivator for the government and commercial enterprises globally. The agency can further improve its delivery and efficiency by using autonomous mail delivery vehicles and further monetizing its geolocation data. Using robotics for the last-mile postal delivery will help with the efficiency and speed of sorting and delivering mail and parcels.

On the issue of political interference, Congress has a more significant part to play through legislation. First, the postmaster should not be a presidential appointee. Still, laws are adjusted to have them be employed by a team of experts in the industry as that would limit sitting presidents from appointing their cronies who may be incompetent and bring down the agency. Human resource management also requires improvement is the core of the agency's operations. Rigorous interviews are required when hiring the various employees to ensure that only the best and most talented gets to the different position. Offering better time in software, competitive

salaries, and good working conditions can help attract talented individuals to fill the other positions.

Conclusion

This case study concludes that the postal services in America can achieve a positive outcome by investing in information technology and related advancements. The benefits of technology in the postal service outweighs the risks and concern going forward. Technology is becoming an essential driver for the operations at USPS. Adopting more digital solutions will enable the postal service opportunity to explore areas beyond the core services it offers. The agency also requires the intervention of Congress to allow the exploration of more areas of businesses to make them more profitable.

References

Buko, J., Bulsa, M., & Makowski, A. (2022). Spatial Premises and Key Conditions for the Use of UAVs for Delivery of Items on the Polish Courier and Postal Services Market. *Energies*, *15*(4), 1403.

Finger, M., & Bukovc, B. (Eds.). (2014). *Postal services in the digital age* (Vol. 6). IOS Press.

Fuglsang, L., & Møller, J. K. (2014). 6 Framing Innovation in Postal Services. *Framing innovation in public service sectors*, 112.

Hillebrand, A., Thiele, S., Junk, P., Hildebrandt, C., Needham, P., & Kortüm, M. (2016). Technology and change in postal services–impacts on consumers. *Bad Honnef: WIK Consult and ITA Consulting*.

Isa, S. M., & Kiumarsi, S. (2019). The impact of service quality in postal services: The mediating role of self-service technology. *International Journal of Services and Operations Management*, *33*(3), 395-419.

Kamarck, E. (2009). The Future of the United States Postal Service.

Marr, B. (2019). *Artificial intelligence in practice: how 50 successful companies used AI and machine learning to solve problems*. John Wiley & Sons.

Trimble, J. (2016). Postal Service Development Requires ICT and Big Data. In *Plenary Presentation for ADMI National Conference*.

Healthcare Analytics and Services Management

ABSTRACT:

This study explores the relationship between Communication Quality, Performance Score, and a patient's Recommendation of a Hospital by using Importance-Performance Map Analysis. Value-Based Purchasing (VBP) and Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) scores were used for the analysis. Study results indicate that Care Transition has a significant impact on a patient's inclination to provide positive recommendation for the hospital where care was received.

INTRODUCTION

Health institutions are taking great care to increase patient satisfaction. Studies such as the comfort of the waiting room, a comfortable delivery room, free Wi-Fi service, TV and refrigerator in each room, quality of meals will positively affect the evaluation of the patient. Today, patient satisfaction has become so important that it affects performance scores and reimbursement, as well as the preferences of patients who choose their hospitals based on hospital ratings. If a patient is satisfied with the hospital where he/she is receiving treatment, it is highly that the patient will prefer the same hospital in the future, and s/he will recommend this hospital to his/her relatives.

There are many factors when choosing a hospital for patients. For some patients, the physical features of the hospital, the condition of the treated room and utilities in it, and even the location of the hospital take the first place. For some patients, the names of doctors are in the first place on the list of importance, how many years of experience they have, and even from which medical school they graduated. In this study, we investigated the effect of communication with the patient on patient satisfaction and the effect of this on recommending the hospital to others.

The patient-centered care (PCC) approach is not only providing medical assistance to the patient, but also investigating the reasons for visiting the doctor, meeting the emotional needs of the patient, and giving them the opportunity to choose between different treatment options.

Patients may react very differently to the same disease. Therefore, the effect of treatment methods on each patient will not be the same. Deciding the treatment process together with the patient, taking into consideration his wishes and needs, gives the patient confidence.

CMS recently launched the Value-Based Purchasing (VBP) program, where healthcare organizations and providers are reimbursed based on the quality of care rather than the number of services. One of the factors affecting the quality of care was determined as patient satisfaction. CMS uses the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey to evaluate patients' hospital experience and satisfaction.

Apart from the different questionnaires that were previously conducted in local or specific hospitals, the HCAHPS survey is a national, standardized and publicly reported survey[1]. Patients receive this survey after discharge in many ways such as via mail, phone call, or in follow-up appointment. In the survey patients are asked about demographic items, screening and their experiences during their time in hospital. The questions of patient's experience are asked under 10 major topics: nurse communication, doctor communication, responsiveness of hospital staff, pain management, communication about medicines, discharge information, cleanliness, quietness, willingness to recommend the hospital and overall rating of the hospital. Then, the HCAHPS scores are calculated for each hospital and reported quarterly in a year.

Nurse Communication

The fact that nurses play an active role in the patient-centered treatment results in improved patient outcomes [2]. Manary et.al. claim that when they analyze patient experience scores, communication with nurses is more predictive than other factors [3]. In patient-centered care, many communication methods are used to meet the expectations of the patients and increase their satisfaction [2]. It was determined that the cue-responding behavior of nurses was appreciated by the patients [4].

Communication with nurses is one of the strongest predictors of dominant influence on the overall satisfaction of the patient from hospital experience [5].

Doctor Communication

Little et.al., in their observational study, found that patients will be less satisfied if doctors do not offer a positive, patient-centered approach, and may have higher referral rates [6]. The importance of communication with doctors may differ in patients suffering from different diseases. Especially in serious diseases such as cancer treatment, while the importance of other parameters such as the comfort of the room decreases, the importance of the treatment to be applied and the physician who will decide on this treatment increase even more.

Moreover, patients who believe that their doctor or the nurses who take care of him will help him get better health fight their illness more easily. For instance, in the study [7] conducted on patients with chronic back pain, it was stated by patients that open communication was the most important parameter in the patient-doctor relationship .

Communication about Medicine

In the HCAPS survey, one of the factors affecting patient satisfaction scores is Communication about Medicine (ComMed). In the questionnaire, questions such as whether the medicine given to the patients is explained and whether the side effects are explained or not are asked. The study [8] shows that approximately 50% of patients admitted in hospital or ambulatory care patients or the patients in nursing homes receive 1 or more unnecessary drugs.

Taking unnecessary medications can adversely affect patients' existing illnesses, and may result in hospitalization or even death especially in older patients[9]. Cals et. al. [10] shows that unnecessary prescription of antibiotics in lower respiratory tract infections decreases with good communication skills. Explaining the use of drugs to patients should be answers to these questions: Why the patient should take this medicine? How will the patient take the medicine and in what dosage? How will this medicine affect the patient? How long should the medicine be taken? What are the potential side effects? In which case should you inform your doctor? Although Communication about Medicine is directly affecting patient satisfaction, this procedure is not done as it should be in some hospitals. For instance, the study [11] showed using CMS's Hospital Compare Data that rural hospitals is better at "communication about medicines" than urban hospitals.

Care Transition

Care transitions in health settings are defined as the transition of the patient from the current location to another station in the same environment or to another health environment [12]. During the transfer, all precautions must be taken for the health of the patient, and the health practitioners who will accompany the transfer must be trained in this regard. Since patients are already in a vulnerable state, they may feel dysfunction, dizziness or pain, as well as risk of infection. When patients arrive at receiving health setting, often the problems encountered are the lack of information about the service provided and medications prescribed by the sending healthcare provider[13].In addition, the detailed care plan of the patient after

discharge, the education of the patient's family and the caregiver, and their communication with their doctors in this process are part of care transition.

Responsiveness of hospital staff

This topic is related with how responsive the hospital staff helped when the patients used call button. In “The person and community engagement domain”, which is the part of Total Performance Score [14], the responsiveness of hospital staff is shown as a percentage of patients who report that staff respond “Always” when they need help. The timely help of hospital staff is especially important for patients who cannot meet their daily needs such as getting to the bathroom or eating meal. While friendly and helpful approach pleases patients, inadequate respect and inadequate empathy leave negative traces on patients [15], as well as affecting hospital evaluations of patients negatively.

Discharge Information

The doctor should talk to the patient about the issues that should be considered after treatment, and if there are symptoms and possible risks, these should be given to the patient in written information. In a study of 377 patients discharged from an academic medical center, it was observed that 46.3% of the patients received their discharge information on the day of discharge, but 24.7% of the patients received it after one week[16].

With the correct and complete information given to the patient, the unnecessary and costly re-hospitalization can be prevented [17], [18]. The study on the effect of discharge information on patients after surgery shows that the patients who received sufficient and timely discharged information seeks assistance, calls or accesses to hospital less than the patients who received insufficient information or no information [19].

Recommend Hospital

In the past, people shared their hospital experience only with their closest relatives or friends. Other people would have a good or bad idea about that hospital with what they heard. However, with the introduction of the internet into our lives, ideas were shared on many platforms, and comments were made about hospitals, doctors, nurses. Thus, instead of the information heard from a few people, there was an opportunity to see tens or even hundreds of people's comments, likes, and complaints and compare them. As a result, many people began to choose their hospital or doctor after seeing general reviews about it. From a business perspective, hospitals and doctors were forced to leave a positive impression on the patient.

Providing patient satisfaction is the first condition of recommending a hospital. However, patients who say they are satisfied with the hospital may not be equally inclined to recommend the hospital [20]. Because patients who report satisfaction for their own standards may feel an extra responsibility when recommending. A parameter that the patient can tolerate when evaluating the hospital, such as bad nurse attitudes, thinks that he can make the person he recommends unhappy. If the person they recommend is not satisfied with the hospital, they will think that they mislead that person. Therefore, although patient satisfaction and the tendency to recommend should be close, they should be evaluated differently.

RESEARCH METHODOLOGY

Analysis of Structural Model

After running the PLS algorithm and re-sampling procedure called Bootstrapping to re-confirm the results (5000 random samples) for the structural model, the output reveals the model estimates as shown in Figure 2. In this process, as our model is formative, Cronbach's alpha values, composite reliability values and average variance extracted values were found 1.00. Outer weights and outer VIF values are presented in Table 1. As shown in Table 1, all VIF values are considerably below the threshold of 5, providing support that collinearity is not critical issue.

Table 1. Outer Weights

Outer Weights	Mean	Std.Dev.	P values	Outher VIF values
CareTr → Communi	0,576	0,019	0,000	3,077
ComA → Communi	0,034	0,018	0,055	2,533
Discln → Communi	0,056	0,016	0,001	2,020
Doctor → Communi	0,130	0,018	0,000	2,475
Overall ← Overall rating star of hospital	1,000	0,000		1,000
RecHos ← Recommendation of hospital	1,000	0,000		1,000
Staff → Communi	0,036	0,018	0,045	2,719
Nurse → Communi	0,286	0,021	0,000	3,182

For *discriminant validity*, Fornell-Larcker criterion [21] was applied. According to this criterion, the square root of AVE values of all the reflective constructs should be higher than

the inter-construct correlations. This situation indicates the discriminant validity. In our model, the results showed that the discriminate validity was acceptable as shown in Table 2. In the table, bold diagonal values show squared root of AVE values, and off-diagonal values represent the correlations of each construct with other constructs.

Table 2: Discriminant Validity-Fornell*Larcker Criterion

	Communication	Overall star rating of hospital	Recommendation of hospital
Overall star rating of hospital	**0,834**	1,000	
Recommendation of hospital	0.799	**0,859**	1,000

Construct validity is achieved when fitness indexes for a construct are the required level. The values of fitness indexes were found required level (Table 3).

Table 3: Fit Summary

	Saturated model	Estimated model
SRMR	0.01	0.03
d_ULS	0,04	0.04
d_G	0.00	0.10
Chi-Square	154,559	1,540.631
NFI	0.99	0.93

R^2 values of the latent endogenous variable were found as 0.70 for overall star rating of hospital value and 0.64 for recommendation of hospital. The former indicates that 70 % of the overall star rating of hospital value could be estimated by the communication (exogenous construct) into the model. At the same time, 64 % of the recommendation of hospital value can be measured by using the communication construct.

Table 4. Results of Structural Analysis of the Model

	Path coefficients	Standard deviation	T statistics	P- values
Communication → Overall star rating of hospital	0,834	0,005	155,021	0.000**
Communication → Recommendation of hospital	0, 799	0,006	125,974	0.000**

Note: ** 0.01; * 0.05

As shown in Figure 2 and Table 4, the results indicated that the coefficient of path from communication to overall star rating of hospital ($\beta = 0.84$, $p < 0.01$) and, the coefficient of path from communication to recommendation of hospital ($\beta = 0.80$, $p < 0.01$) were statistically found as significant and positive. Thus, H_1, H_2 hypotheses were supported.

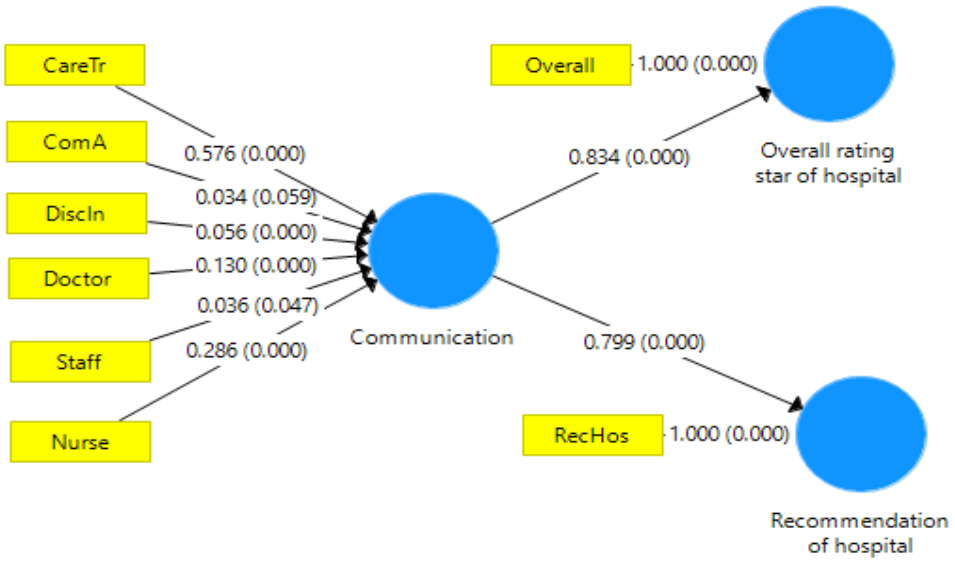


Figure 2: Results of the Structural Model Analysis-Path Coefficients

Importance-performance map analysis

We extended the standard PLS-SEM results reporting path coefficient estimates and other parameters by adding a procedure that considers the average values of the latent variables scores [22]. "IPMA represents the predecessor constructs' importance in predicting a specific target construct, with their average latent variables scores indicating their performance" (Hair et.al., 2018, p.105). Identifying which are predecessor constructs not only have relatively high importance for predicting target construct (those that have a strong total effect) but also have relatively low performance (low average latent variables scores) (Ringle and Sarstedt, 2016). Therefore, the fields that have high importance, but low performance can be determined and suggested many improvements to practitioners and policy makers. Hair et.al (2018) suggested to examine only the predecessor construct which has the largest effect on target construct, although they stated that it's not necessary, instead of examining all of predecessor constructs in the model (p.109). Because we have one latent construct but two target constructs in our model, we separately applied IPMA for two target constructs. Before the IPMA analysis was applied, whether three requirements for analysis meet or not in the PLS path model had been checked. First, all indicators were measured by equidistance scale (e.g. on 5-point scale). So that, transforming of the 5-point scale into the on 0-100 point scale is easily realized. Second, all the indicators coding has the same scale direction. Third, the outer weights estimates are positive. Because the predecessor construct (communication) has high effect on both the overall star rating score of hospital ($\beta = 0, 83$) and the recommendation of hospital ($\beta = 0, 80$) as two target constructs, it is particularly advantageous to further analyze this predecessor construct on the indicator level. "In this case, the indicators describe aspects that shape the corresponding construct, while their weights indicate each aspect's importance in this respect. Therefore, aspects underlying indicators with high weights should be given more attention to identify managerial or policy actions aimed at improving the target construct's performance" (Hair et.al., 2018, p.109). In our model, CareTran, ComAboMed, Disclnfo, DoctorCommunication, StaffCommunication, NurseCommunication variables were determined as indicators of communication quality.

At this point, after computing the importance and performance values, all information required to draw the importance-performance map is available. In the first step, we selected the overall star rating of hospital as a key target construct. Then in the second step, we selected the recommendation of hospital as a key target construct. The results of IPMA, for overall star

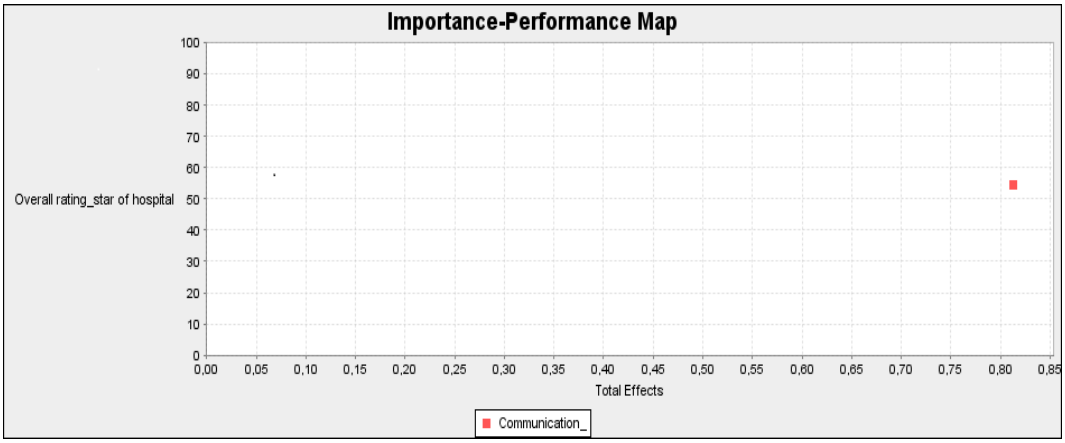
rating score of hospital, on level of latent variable, namely communication, are presented in Table 5 and Figure 3.

Table 5. Construct Total Effect and Performance for Overall Star Rating of Hospital

	Total effect	Performance
Communication	0,81	54,50

In our model, communication variable has a mean value (i.e., performance) of 54,50. This result serve as input for the importance-performance map's performance dimension. A construct's importance in terms of explaining another directly or indirectly linked (target) construct in the structural model is derived from the total effect of the relationship between these two constructs. Total effect is found as 0,81.

Figure 3. Importance-Performance Map Analysis of Communication Quality Effect on Overall Star Rating of Hospital



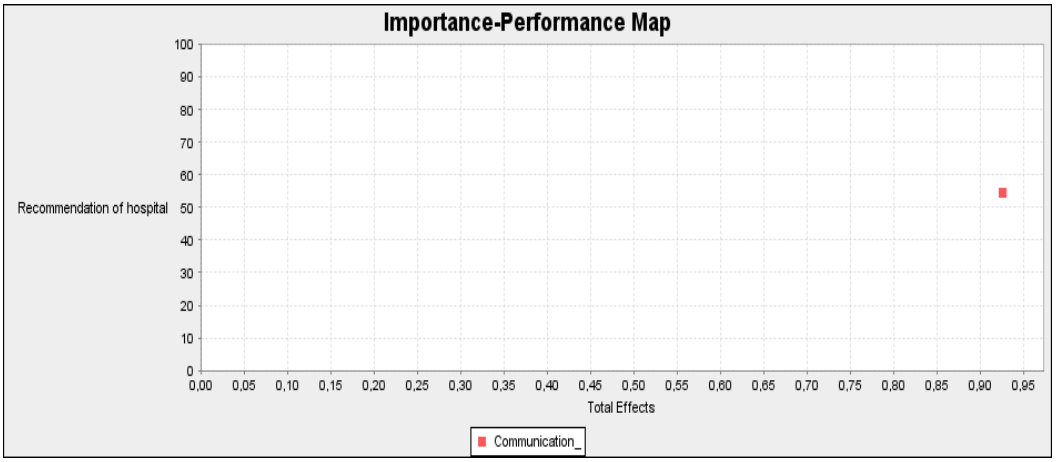
With regard to the recommendation of hospital, the results of IPMA, on level of latent variable, namely communication, are presented in Table 6 and Figure 4. The communication has a mean importance 0,93 and a mean performance of 54,50.

Table 6 Construct Total Effect and Performance for Recommendation

	Total effect	Performance
Communication	0,926	54,496

Table 5 and Table 6 show the values of these maps' importance and performance dimensions, as obtained by the previous IPMA steps. Scatter plotting the information shown in Table 5 and 6 enable us to create an importance-performance map as shown in Figure 3 and Figure 4.

Figure 4. Importance-Performance Map Analysis of Communication Quality Effect on Recommendation of Hospital



The x-axis represents the importance of communication for the target construct, while the y-axis depicts the performance of communication in terms of its average rescaled latent variable scores. For a better orientation, we drawn two additional lines in the importance-performance map: the mean importance value i.e. a vertical line) and the mean performance value (i.e. a horizontal line) of the displayed constructs. These two additional lines divide the importance-performance map into four areas with importance and performance values below and above the average. Generally, when analyzing the importance-performance map, constructs in the lower right area (i.e. above-average importance and below-average performance) represent greatest opportunity to achieve improvement, followed by the upper-right, lower-left, and, finally, the upper left areas. Thereby, the importance-performance map provides guidance for the prioritization of managerial or policy activities of high importance for the aspect underlying the selected target, but which require performance improvement. In our study, Figure 3 shows that communication has a relatively middle performance while it has high importance. Therefore one unit increase in communication performance from 54,50 to 55,50 would increase the performance of overall star rating of hospital by 0,81 points from 70,0 (mean of rescaled overall star rating of hospital variable) to 70,81. On the other hand, Figure 2 shows

that one unit increase in communication performance from 54,50 to 55,50 would increase the performance of recommendation of hospital by 0,926 points from 63,8 (mean of rescaled recommendation of hospital variable) to 64,73.

We can also conduct an IPMA on the indicator level to identify relevant and even more specific areas of improvement. More precisely, we can interpret the rescaled outer weights as an indicator's relative importance compared to that of the other indicators in a specific measurement model. In addition to, we also created separate importance-performance maps for the indicators of each constructs namely recommendation of hospital and overall star rating of hospitals.

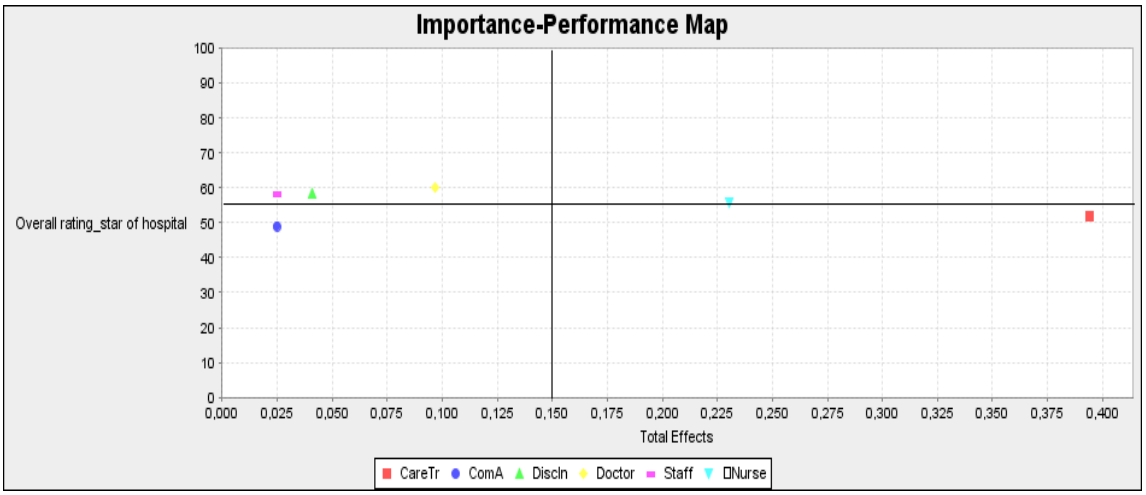
Table 7. Indicator Importance and Performances for Overall Star Rating of Hospitals.

Recommendation	Total Effects	MV Performances
CareTr	0,39	52,09
ComA	0,02	49,00
Discln	0,04	58,40
Doctor	0,09	60,00
Staff	0,02	58,01
Nurse	0,23	55,78
Mean Value	**0,13**	**55,54**

Note: All effects denote unstandardized effects.

The results suggest that indicator Care Transition should be given the highest priority for improvement, since it has the highest relative importance but lower performance relative to Discln, Doctor, Staff and Nurse. A one-unit point increase in Care T 's performance increases the performance of Overall star rating of hospital by Care T's importance value, which is 0,39 (ceteris paribus). Indicators nurse and doctor communication follow with second to third priority (Table 7 and Figure 5).

Figure 5. Importance-Performance Map Analysis of Communication Quality Indicators Effecting on Overall Star Rating of Hospital



Same results are valid for the recommendation of hospitals construct (Table 8 and Figure 6).

Table 8. Indicator Importance and Performances For Recommendation of Hospital.

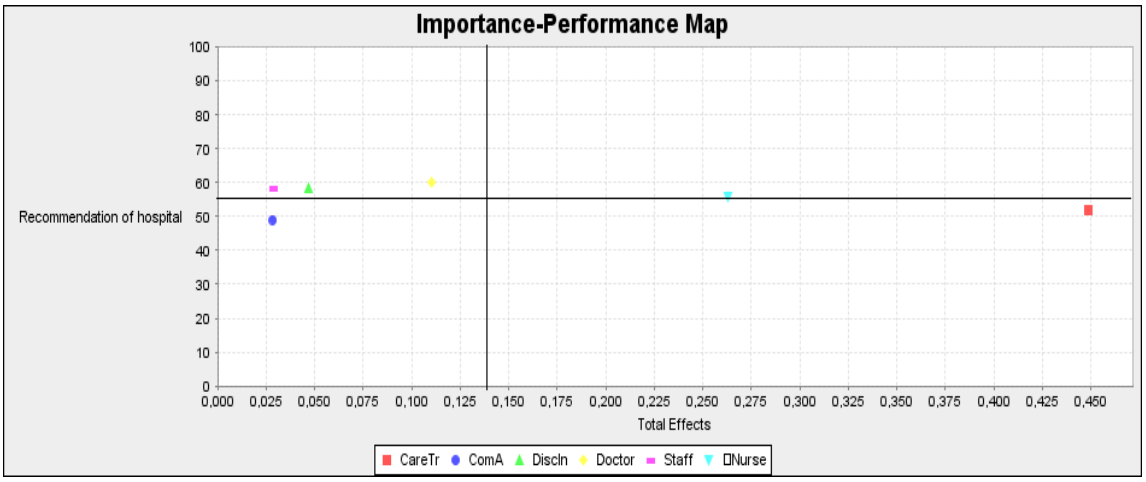
Recommendation	Total Effects	MV Performances
CareTr	0,44	52,09
ComA	0,02	49,00
Discln	0,04	58,40
Doctor	0,11	60,00
Staff	0,02	58,01
Nurse	0,26	55,78
Mean Value	**0,15**	**55,54**

Note: All effects denote unstandardized effects.

The results suggest that indicator Care Transition should be given the highest priority for improvement, since it has the highest relative importance but lower performance relative to Discln, doctor, staff and nurse. A one-unit point increase in Care T ‘s performance increases the performance of Recommendation of hospital by Care T’s importance value, which is 0,44

(ceteris paribus). Indicators nurse and doctor communication follow with second to third priority.

Figure 6. Importance-Performance Map Analysis of Communication Quality Indicators Effecting on Recommendation of Hospital



References

[1] E. Goldstein, M. Farquhar, C. Crofton, C. Darby, and S. Garfinkel, "Measuring hospital care from the patients' perspective: An overview of the CAHPS® Hospital Survey development process," *Health Serv. Res.*, vol. 40, no. 6, pp. 1977–1995, 2005.

[2] C. R. Charlton, K. S. Dearing, J. A. Berry, and M. J. Johnson, "Nurse practitioners' communication styles and their impact on patient outcomes: An integrated literature review," *J. Am. Acad. Nurse Pract.*, vol. 20, no. 7, pp. 382–388, 2008.

[3] M. P. Manary, W. Boulding, R. Staelin, and S. W. Glickman, "The Patient Experience and Health Outcomes," *N. Engl. J. Med.*, vol. 368, no. 3, pp. 201–203, 2013.

[4] R. Uitterhoeve, J. Bensing, E. Dilven, R. Donders, P. DeMulder, and T. Van Achterberg, "Nurse-patient communication in cancer care: Does responding to patient's cues predict patient satisfaction with communication," *Psychooncology.*, vol. 18, no. 10, pp. 1060–1068, 2009.

[5] H. S. Laschinger, L. M. G. Hall, C. Pedersen, and J. Almost, "A psychometric analysis of the patient satisfaction with nursing care quality questionnaire: An actionable approach to measuring patient satisfaction," *J. Nurs. Care Qual.*, vol. 20, no. 3, pp. 220–230, 2005.

[6] P. Little *et al.*, "Observational study of effect of patient centredness and positive approach on outcomes of general practice consultations," *Br. Med. J.*, vol. 323, no. 7318, pp. 908–911, 2001.

[7] E. Schmidt, L. Gramm, and E. Farin, "Communication preferences of patients with chronic back pain in medical rehabilitation," *Schmerz*, vol. 26, no. 1, pp. 69–76, 2012.

[8] J. Tjia, S. J. Velten, C. Parsons, S. Valluri, and B. A. Briesacher, "Studies to Reduce Unnecessary Medication Use in Frail Older Adults : A Systematic Review," *Drugs Aging*, vol. 30, pp. 285–307, 2013.

[9] I. A. Scott *et al.*, "Reducing Inappropriate Polypharmacy The Process of Deprescribing," *JAMA Intern. Med.*, vol. 175, no. 5, pp. 827–834, 2015.

[10] J. W. L. Cals, C. C. Butler, R. M. Hopstaken, K. Hood, and G.-J. Dinant, "Effect of point of care testing for C reactive protein and training in communication skills on antibiotic use in lower respiratory tract infections: cluster randomised trial," *BMJ*, vol. 338, no. b1374, pp. 1–10, 2009.

[11] M. M. Casey and G. Davidson, "Patient Assessments and Quality of Care in Rural Hospitals," 2010.

[12] E. A. Coleman, "Falling through the cracks: Challenges and opportunities for improving," *J. Am. Geriatr. Soc.*, vol. 51, no. 4, pp. 549–555, 2003.

[13] M. A. Anderson and L. B. Helms, "Communication between continuing care organizations," *Res. Nurs. Health*, vol. 18, no. 1, pp. 49–57, 1995.

[14] "Total Performance Score." [Online]. Available: https://www.medicare.gov/hospitalcompare/data/total-performance-scores.html. [Accessed: 06-Feb-2020].

[15] E. Jangland, L. Gunningberg, and M. Carlsson, "Patients' and relatives' complaints about encounters and communication in health care: Evidence for quality improvement," *Patient Educ. Couns.*, vol. 75, no. 2, pp. 199–204, 2009.

[16] L. I. Horwitz *et al.*, "Comprehensive quality of discharge summaries at an academic medical center," *J. Hosp. Med.*, vol. 8, no. 8, pp. 436–443, 2013.

[17] R. L. Evans and R. D. Hendricks, "Evaluating hospital discharge planning: a randomized clinical trial," *Med. Care*, vol. 31, no. 4, pp. 358–370, 1993.

[18] J. K. Magilvy and J. M. Lakomy, "Transitions of Older Adults to Home Care," *Home Health Care Serv. Q.*, vol. 12, no. 4, pp. 59–70, 1991.

[19] A. Henderson and W. Zernike, "A study of the impact of discharge information for surgical patients," *J. Adv. Nurs.*, vol. 35, no. 3, pp. 435–441, 2001.

[20] S. H. Cheng, M. C. Yang, and T. L. Chiang, "Patient satisfaction with and recommendation of a hospital: Effects of interpersonal and technical aspects of hospital care," *Int. J. Qual. Heal. Care*, vol. 15, no. 4, pp. 345–355, 2003.

[21] C. Fornell and D. F. Larcker, "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *J. Mark. Res.*, vol. 18, no. 1, pp. 39–50, 1981.

[22] C. Fornell, M. D. Johnson, E. W. Anderson, J. Cha, and B. E. Bryant, "The American Customer Satisfaction Index: Nature, purpose, and findings," *J. Mark.*, vol. 60, no. 4, pp. 7–18, 1996.

High Cost Procedures in US Hospitals: Study of Influencing Factors

Ravi Chinta, Ph.D.
Associate Dean and Professor, Strategic Management
School of Business and Public Administration
University of District of Columbia
4200 Connecticut Ave NW
Washington, DC 20008
Tel: (513) 262-5512
ravi.chinta@udc.edu

Keywords: Hospital Charges; in-patient care; High-cost procedures; Hospital Location; Patient demographics; AHRQ-NIS Data

High Cost Procedures in US Hospitals: Study of Influencing Factors

Abstract

It is well known that hospitalization costs in the US vary widely. The significant and apparent variance in hospital charges and inpatient care in the U.S. has perplexed the general public including many stakeholders such as the healthcare regulators and insurers. While the clinical side of inpatient care has been undergoing tremendous progress and standardization, the overall cost of healthcare has been ballooning. The purpose of this research is to conduct statistical analyses that reveal the sources of variance in hospital charges and inpatient care using the annual data from the AHRQ's (Agency for Healthcare Research and Quality) HCUP's (Hospital Cost and Utilization Project) NIS (National Inpatient Sample) database. The focus is on non-clinical factors such as patient age, gender, income and race and hospital location data as independent variables to explain variance in hospital charges and inpatient care. The research sample is the liver transplant cases in 2019 sampled in the NIS 2019 database. The regression results show patient age and gender as well as payer affect the number of diagnoses; and hospital charges are affected by age, payer and hospital location. Implications suggest that there is more room for standardization of the number of diagnoses and procedures across regions in the US. Results also reveal that race and income do not have any effect on hospital charges and inpatient care. The study contributes to an understanding of non-clinical factors in the explanation of variance in hospital charges and inpatient care.

Introduction

There are many high-cost procedures in US hospitals. Liver transplant is chosen for the study because it is one of the very high cost procedures in hospitals. Orthotopic Liver Transplantation (OLT) has proven to be very effective as treatment for end stage liver disease. Despite its effectiveness, its high cost has received much attention from both the academic and practitioner community. According to Organ Procurement and Transplantation Network, more than 9000 liver transplantations across the United States have been performed in 2021 alone across the United States (*National Data - OPTN*, n.d.).

Literature review

Liver transplants are so expensive that in some states in the US (e.g., Oregon), Medicare does not cover the procedure. The current study explores factors that can explain these cost disparities. It uses a representative 2019 dataset and limits its attention to the underlying demographic factors and hospital characteristics. Past research on cost variances has explained cost variances from different perspectives. A previous study developed a model using twelve variables and found that the model explained 37% of the total variation in charges (Brown et al., 1998). Another study examined the relationship between the total cost incurred by liver transplantation (LT) recipients and it used the Model for End-Stage Liver Disease (MELD) score at the time of transplant as a possible factor explaining hospital costs. This study used a database linking billing claims from a large private payer with the Organ Procurement and Transplantation Network registry and found no significant differences in post-transplant hospital charges but a difference in costs during pre-transplant and during hospital stays at the time of transplant (Buchanan et al., 2009). Another study found creatinine and blood transfusion were significantly associated with prolonged Length of Stay (LOS), ICU LOS, and days on the ventilator and concluded that elevated serum creatinine and blood transfusions are the most critical determinants of increased resource utilization and hospital expenditure in LT (Ruiz et al., 2018). Yet another finding concludes that older recipients, who had alcoholic liver disease or were severely ill were the most expensive to treat, suggesting that organ allocation criteria may affect

transplant costs (Showstack et al., 1999)). This research makes two significant departures from many previous studies: First, it avoids linking databases from multiple sources that might introduce disparities in using data due to data collection and labeling inconsistencies. Second, it uses a national stratified random sample rather than from a single source that might not represent the true national costs. Third, we limit ourselves to structural (non-clinical) rather than clinical factors influencing costs. The study does not ignore the clinical factors but these are reflected in its research model through structural equivalents such as the number of in-patient diagnoses and treatments.

This research is prompted by a need to explain the wide variance in hospital charges and in-patient care (length of stay in the hospital, # of diagnoses and number of procedures) that is evidenced in liver transplant cases across the United States. AHRQ's mission is to facilitate research in healthcare costs and processes. Without going into the complexities of clinical decision making, we wish to examine the AHRQ data to empirically reveal the impact of non-clinical variables in AHRQ data set on the hospital charges and in-patient care (length of stay in the hospital, # of diagnoses and number of procedures). Our research focus is thus limited by the data elements in the AHRQ data set, and HIPAA requirements prohibit us from linking this data with other data sets available elsewhere.

Grounded Research Model and Hypotheses

The AHRQ data set for 2019 in the US shows that for a sample of 450 liver transplants, the hospital charges ranged from a minimum of $157 to a maximum of $1.15 million with a mean value of $364,000. For the same sample data, the number of diagnoses ranged from 3 to 39 with an average of 17; the number of treatments ranged from 1 to 19 with an average of 3.5 and the length of stay in hospital ranged from 1 to 45 with an average stay of 9 days.

As noted earlier, we are limited to using only the patient data within the AHRQ data set. Our aim is to explain the variance in hospital charges and in-patient care by using the non-clinical variables in the data set. Patient age, gender, income and race are the patient specific variables that will be used in our analysis. Among these, only age is a continuous variable and the other three are categorical variables. Hospital division is also a categorical variable.

Research sample and variables

The Agency for Healthcare Research and Quality (AHRQ) is one of twelve agencies within the United States Department of Health and Human Services. AHRQ currently has an annual budget of more than $488.8 million (FY2022) to compile open government data for healthcare research. Since the early 1990s, AHRQ's "The Health Care Cost and Utilization Project (HCUP)" has been collecting data from 4,568 hospitals which is a representative sample of hospitals across the U.S. The unit of analysis in the HCUP databases represents a single inpatient episode, from hospital admissions to discharge. Records from VA hospitals, hospitals on Indian Reservations, and long-term care hospitals were excluded from our study. The hospitals employ a DRG (Diagnosis Related Group) code from 000 to 999 to classify each admission. In a given year, the sample consists of more than 7 million records with information for each admission on about 250 variables.

The HIPAA Privacy Rule sets national standards for patient rights with respect to health information. This rule protects individually identifiable health information by establishing conditions for its use and disclosure by covered entities. The HCUP databases conform to the definition of a limited data set. A limited data set is healthcare data in which 16 direct identifiers, specified in the Privacy Rule, have been removed. Under HIPAA, review by an institutional review board (IRB) is not required for use of limited data sets.

We focused on readmissions coded as DRG=6 (Liver Transplants) for the year 2019. Our objective was to empirically understand if non-clinical factors such as the patient demographic variables (race, income, age, and gender) impact hospitalization charges and in-patient care. We did not question the clinical decisions made by the doctors who diagnose and treat patients with diligent care providing highly patient-specific care (number of diagnoses, number of treatments, length of stay in hospital, etc.) aiming for best patient outcomes. Our aim is to empirically examine the non-clinical predictors and implications of hospital charges and in-patient care.

Different DRGs have different clinical protocols for patient care. This is the main reason why we focused only on liver transplants for our study. As would be expected, even within liver transplants, there is considerable variation in patient cases such as severity of illness, risk of morbidity, age, sex, etc. that we need to consider in our study. Likewise, there is a hospital vector of many hospital specific variables such as size, location, type of hospital, etc.

Research Sample

Each record in the NIS database represents one inpatient episode. All our study variables are from the NIS 2019 database in the HCUP (https://www.hcup-us.ahrq.gov/db/nation/nis/nisdde.jsp). The total number of patient records across all DRG codes in 2019 was 7,083,805 which had been collected from a stratified sample of 4,568 hospitals in the U.S. The hospital composition was 13% from the Northeast, 30% from the Midwest, 38% from the South, and 19% from the Western region. The hospitals in the sample were 20% government, non-federal hospitals; 64% private, not-for-profit hospitals, and 16% private, investor-owned hospitals. Of the total 7,083,805 records from 4,568 hospitals, the number of records in the database for liver transplants (DRG code = 6) was 450 discharge records, which is the sample size for our research study. Table 1 presents the univariate statistics of our sample of 450. The variable descriptions are detailed at the AHRQ web site (https://www.hcup-us.ahrq.gov/db/nation/nis/nisdde.jsp).

Data analysis and results

The research model in Figure 1 above indicates that the hospital charges and in-patient care (length of stay in the hospital, # of diagnoses and number of procedures) are the dependent variables. The independent variables are patient age, gende, income, race and hospital division. Thus, four regression models are used to examine the relationships between the dependent and independent variables. The results of the regression models are shown in Table 1 below.

Table 1: Regression Results				
	Model 1	Model 2	Model 3	Model 4
	Dependent Variable in Regression Models 1-4			
Predictor Variable (Measure)*	**Hospital Charges**	**# of Diagnoses**	**# of Procedures**	**Length of Stay**
Age (Continuous)	F=10.8, Sig (.001)*	F=9.5, Sig (.002)*	F=0.1, Not Sig.	F= 19.9, Sig (.000)*
Gender (Categorical)	F=0.2, Not Sig.	F=3.4, Sig (.068)*	F=2.0, Not Sig.	F= 6.0, Sig (.015)*
Income (Categorical)	F=1.3, Not Sig.	F=0.6, Not Sig.	F=0.6, Not Sig.	F=1.1, Not Sig.
Race (Categorical)	F=1.0, Not Sig.	F=1.0, Not Sig.	F=0.5, Not Sig.	F=0.9, Not Sig.
Payer (Categorical)	F=2.1, Sig (.06)*	F=3.3, Sig (.006)*	F=0.4, Not Sig.	F=1.3, Not Sig.
Hospital Division (Categorical)	F=15.9, Sig (.000)*	F=1.4, Not Sig.	F=4.0, Sig (.000)*	F=2.7, Sig (.007)*
*Note: Significant results related to the categorical measures have detailed explanations in the paper that show differences across categories relative to a baseline category.				

When an independent variable is a categorical variable, dummy variables have to be created for use in the regression models. If a variable has "N" categories, then (n-1) dummy variables have to be used in the regression model with the excluded category as the reference to which the regression model results have to be compared. If the independent variable is a continuous variable, then it can directly be used in the regression model. Therefore, the regression results are discussed next taking each independent variable at a time.

Age is a continuous variable with a mean value of 54 years, modal value of 57 years and median value of 58 years in the sample of 450 liver transplants. The regression of hospital charges with age as the independent variable shows a significant negative beta coefficient (F=10.8, p=0.001). Interpretation of the slope of the regression means that the baseline hospital costs are $447,756 and with each increment of 1 year in age there would be a reduction of $1,550. Thus, for a 60 year old liver transplant patient, the predicted hospital charge would be $354,756. Age was a statistically significant (F=9.5, p=002) determinant of number of diagnoses with a baseline of 14 diagnoses, but was not a statistically significant determinant of number of procedures. Finally, age was a statistically significant (F=19.9, p=000) determinant of length of stay with a baseline of 12 days.

Income is a categorical variable measured in the AHRQ data set as in four quartiles (0-25 percentile), (26-50 percentile), (51-75% percentile) and (76-100% percentile) of average income of the ZIP code of the patient is coming from. The sample of 450 patients is 23% in 1st quartile, 25% in 2nd quartile, 27% in 3rd quartile, 22% in the 4th quartile, and missing income data is 3%. The regression of hospital charges with income as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as the 1st quartile in interpreting the regression results. The results in Table 2 show there were no statistically significant differences among income categories in all the four regression models. Income did not impact hospital charges or in-patient care variables (# of diagnoses, # of procedures and length of stay).

Gender is a categorical variable with two categories (Male and Female). The sample of 450 patients is 65% male and 35% female. The regression of hospital charges with gender as the independent variable shows no statistically significant relationship. However, gender was a statistically significant (F=3.4, p=068) determinant of number of diagnoses with a baseline of 17 diagnoses for the males and an incremental increase of 1 diagnosis for the female group. Gender was not a statistically significant determinant of the number of procedures. Finally, gender was a statistically significant (F=6.0, p=015) determinant of length of stay with a baseline of 8.6 days for the male group and an incremental increase of 1.1 days in length of stay for the female group.

Race is a categorical variable measured in the AHRQ data set as in 6 categories. The sample of 450 patients is 73.5% White, 7% Black, 14% Hispanic, 3% Asian, 0.5% Native American and 2% Other. The regression of hospital charges with race as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as White in interpreting the regression results. The results in Table 2 show there were no statistically significant differences among race categories in all the four regression models. Race did not impact hospital charges or in-patient care variables (# of diagnoses, # of procedures and length of stay).

Payer is a categorical variable measured in the AHRQ data set as in 6 categories. The sample of 450 patients is 32% Medicare, 14% Medicaid, 48% Private Insurance, 0.7% Self-pay, 0.9% No Charge and 4.4% Other. The regression of hospital charges with Payer as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as Medicare in interpreting the regression results. The results in Table 2 show the Hospital Charges for the baseline Medicare category was $354,220. Compared to this baseline the Medicaid category had a statistically significant higher incremental charge of $54,983, and similarly the Self Pay group patients relative to the Medicare group were being charged a higher incremental charge of $168,385. The other 3 categories of Payer showed no statistically significant differences in Hospital Charges relative to the Medicare group.

With regard to # of diagnoses, the results in Table 2 show the # of Diagnoses for the baseline Medicare category was 18 diagnoses. Compared to this baseline the Medicaid category had a statistically significant 2.7 diagnoses lower, and similarly the Private Insurance group patients relative to the Medicare group had 1.7 diagnoses lower, and the No Charge group patients had 7.3 diagnoses lower compared to the Medicare group. The other 2 categories of Payer showed no statistically significant differences in # of Diagnoses relative to the Medicare group.

Payer variable did not impact # of procedures and length of stay. That is, all Payer categories had no differences in # of procedures and length of stay.

Hospital Division is a categorical variable measured in the AHRQ data set as in 9 categories. The sample of 450 patients is 3% New England (Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut), 13% Middle Atlantic (New York, Pennsylvania, New Jersey), 14% East North Central (Wisconsin, Michigan, Illinois, Indiana, Ohio), 7% West North Central (Missouri, North Dakota, South Dakota, Nebraska, Kansas, Minnesota, Iowa), 22.2% South Atlantic (Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida), 6.4% East South Central (Kentucky, Tennessee, Mississippi, Alabama), 16.4% West South Central (Oklahoma, Texas, Arkansas, Louisiana), 8% Mountain (Idaho, Montana, Wyoming, Nevada, Utah, Colorado, Arizona, New Mexico) and 10% Pacific (Alaska, Washington, Oregon, California, Hawaii). We do not have the more granular state level data.

The regression of hospital charges with Hospital Division as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as New England in interpreting the regression results. The results in Table 2 show the Hospital

Charges for the baseline New England category was $293,742. Compared to this New England baseline, four other groups showed statistically significant higher incremental charges as follows - Mid Atlantic (higher incremental charge of $180,028); East North Central (higher incremental charge of $80,389); Mountain (higher incremental charge of $100,304); and Pacific (higher incremental charge of $224, 117) The other 2 categories of Hospital Division (viz., East South Central and West South Central) showed no statistically significant differences in Hospital Charges relative to the New England group.

Hospital Division variable did not impact # of Diagnoses. That is, all Hospital Division categories had no differences in # of Diagnoses. The baseline l# of Diagnoses for the New England group was 15.4 diagnoses.

With regard to # of procedures, the results in Table 2 show the # of procedures for the baseline New England category was 6.4 procedures. Compared to this baseline the New England all other categories of Hospital Division showed statistically significant lower number of procedures ranging from 2.2 to 3.5 lower relative to New England. .

The Hospital Division variable did not impact length of stay. That is, all Hospital Division categories had no differences in length of stay. The baseline length of stay for the New England group was 9.3 days. Next we discuss the above results and link them to existing literature with implications for addressing the widely observed variance in hospital charges and in-patient care.

Discussion, Implications and Limitations

Age has been studied in many earlier studies as a key determinant of hospital charges and in-patient care for a wide variety of hospitalizations. Past research found evidence attesting that healthcare costs increase with individuals' age (Jacobzone, 2003; Peters, 2006; Farooqui & Farooqui, 2009). Our findings are consistent with past research on age and health. However, our results show that for liver transplants the # procedures are unaffected by the age variable. Our finding is also corroborated by studies that found non-clinical factors other than age might also contribute to hospital cost (Seshamani & Gray, 2004).

Income variable in our study is a crude and aggregate measure based on the ZIP code of the patient and plugging that ZIP code in one of the four quartiles of national income. Hence we do not believe that our findings, which show no relationship between income variable and hospital charges and in-patient care, are at best tenuous. Furthermore, while income affects affordability of healthcare, healthcare delivery (# of diagnoses, # of procedures, length of stay and hospital charges) are guided by standardized clinical protocols that are invariant of the income level of the patient.

Gender related findings in our study reveal statistically significant differences only in # of diagnoses and length of stay, and not in # of procedures and hospital charges. Impact of gender in healthcare has been demonstrated in past research studies (Daniel et al., 2018; Manuel, 2018).

Race variable showed no statistically significant relationship with hospital charges or in-patient care. This is a testament to race-blind delivery of care for liver transplants in the US, but is contrary to several other studies that found disparities in access and cost of healthcare based on race (Ruthberg et al., 2020; Bliss et al., 2015; Gulley et al., 2014). Coincidentally, research has shown that access to universal healthcare seems to mitigate racial disparities in access and quality of healthcare (Holtkamp, 2018).

Payer variable reveals interesting results. While the baseline hospital charge for Medicare was $354,220, the hospital charge for "No Charge" was still at $260,125, which means that hospitals are writing off on average $260,125 for each patient in the "No Charge" group. This perhaps explains

why all other Payer categories show hospital charges that are incrementally higher than the reference group (Medicare) charge of $354,220.The hospital charge for “Self-Pay” group was $168,385 higher than the baseline Medicare charge. One implication is that these findings raise an interesting topic for future research to be directed at examining the accounting practices of hospitals to bring to surface the distinctions between charges and costs incurred at the procedural level. Note that Payer variable does not impact the # of procedures or the length of stay.

Hospital Division variable surfaces some interesting regional differences. Our results show that Pacific (Alaska, Washington, Oregon, California, Hawaii) region is the most expensive at more than half million dollars ($517,859) and the least expensive is West South Central (Oklahoma, Texas, Arkansas, Louisiana) at $278,419 on average. One implication is that these findings may provide some broad guidance for medical tourism to locate least cost hospitals for liver transplantation.

As with all research studies, our study suffers from several limitations. One limitation of our study is the lack of an experimental design and the use of only cross-sectional data (Shadish et al., 2002). However, it is morally inadmissible to conduct experiments in high risk procedures such as liver transplants. Another limitation is that we were limited by the variables available in AHRQ’s HCUP database; however, there may be other variables that may affect hospital charges and in-patient care. Another limitation is inherent in the categorical measurement of many of our research variables which limits analysis of variance in the dependent variables using more robust statistical techniques.

Understanding the contribution of non-clinical factors such as hospital location, payer and patient demographics to healthcare costs has several implications for managing the healthcare delivery in the US. While we focused only on liver transplants, replication of our research for other DRG codes in the AHRQ database will yield a comprehensive aggregate picture of where the underlying causes for variance exist. Sharing such empirically revealed insights will then lead to productive healthcare reforms. That is in fact the aim of the Agency for Healthcare Research and Quality which promotes research on healthcare costs and utilization.

This research will also facilitate sharing of best practices from regions that are more cost effective to regions that are less cost effective as revealed by the empirical findings. We recommend AHRQ to organize an annual symposium in Washington DC focused on determinants of variance in hospital charges for various DRG codes. Analytically driven reforms will have a higher chance of successful implementation in the complex healthcare industry.

Conclusion

The study aimed to identify factors influencing high cost procedures in US hospitals. by actual data compiled in hospitals. While increasing codification and standardization of clinical protocols for advanced procedures would decrease the variance in hospital charges and inpatient care, there is still wide and unexplained variance. Our research supports the mission of the medical profession which is to treat all patients well irrespective of their age, gender, income and race. There is a dearth of data-driven research on specific diseases revealing statistically significant non-clinical determinants of hospital charges and in-patient care. Our study addresses this gap. In the early 20th century, Mary Parker Follett emphasized that leaders must have the “ability to grasp a total situation, i.e., see a whole, not a mere kaleidoscope of pieces” (Graham, 1996, p.168). Our research highlights the non clinical variables as part of the whole picture in understanding hospital charges and in-patient care.

References

Bliss, L. A., Yang, C. J., Kent, T. S., Ng, S. C., Critchlow, J. F., & Tseng, J. F. (2015). Appendicitis in the modern era: Universal problem and variable treatment. *Surgical Endoscopy*, 29(7), 1897–1902. https://doi.org/10.1007/s00464-014-3882-2

Brown, R. S., Lake, J. R., Ascher, N. L., Emond, J. C., & Roberts, J. P. (1998). Predictors of the cost of liver transplantation. *Liver Transplantation and Surgery*, *4*(2), 170–176.

Buchanan, P., Dzebisashvili, N., Lentine, K. L., Axelrod, D. A., Schnitzler, M. A., & Salvalaggio, P. R. (2009). Liver transplantation cost in the model for end-stage liver disease era: Looking beyond the transplant admission. *Liver Transplantation*, *15*(10), 1270–1277.

Daniel, H., Erickson, S. M., Bornstein, S. S., Kane, G. C., Gantzer, H. E., Henry, T. L., Lenchus, J. D., Li, J. M., McCandless, B. M., Nalitt, B. R., Viswanathan, L., Murphy, C. J., Azah, A. M., & Marks, L. (2018). Women's Health Policy in the United States: An American College of Physicians Position Paper. *Annals of Internal Medicine*, 168(12), 874–875. https://doi.org/10.7326/M17-3344

Farooqui, T., & Farooqui, A. (2009). Aging: An important factor for the pathogenesis of neurodegenerative diseases. *Mechanisms of Aging and Development*, 130 (4), 203–215. https://doi.org/10.1016/j.mad.2008.11.006

Graham, P. Ed. (1996). *Mary Parker Follett: Prophet of Management.* Boston, MA: Harvard Business School Press.

Gulley, S. P., Rasch, E. K., & Chan, L. (2014). Difference, disparity, and disability: A comparison of health, insurance coverage, and health service use on the basis of race/ethnicity among US adults with disabilities, 2006–2008. *Medical Care*, 52(Supplement 3), S9–S16. https://doi.org/10.1097/MLR.0000000000000129

Holtkamp, M. D. (2018). Does race matter in universal healthcare? Stroke cost and outcomes in US military health care. *Ethnicity & Health*, 25(6), 888–896.

Jacobzone, S. (2003). Ageing and the challenges of new technologies: Can OECD social and healthcare systems provide for the future? *The Geneva Papers on Risk and Insurance Issues and Practice*, 28(2), 254–274. https://doi.org/10.1111/1468-0440.00222

Manuel, J. (2018). Racial/ethnic and gender disparities in health care use and access. *Health Services Research*, 53(3), 1407–1429. https://doi.org/10.1111/1475-6773.12705

National data—OPTN. (n.d.). Retrieved June 24, 2022, from https://optn.transplant.hrsa.gov/data/view-data-reports/national-data/#

Peters, R. (2006). Ageing and the brain. *Postgraduate Medical Journal*, 82(964), 84–88. https://doi.org/10.1136/pgmj.2005.036665

Ruiz, J., Dugan, A., Davenport, D. L., & Gedaly, R. (2018). Blood transfusion is a critical determinant of resource utilization and total hospital cost in liver transplantation. *Clinical Transplantation*, *32*(2), e13164.

Ruthberg, J., S; Khan, H. A., Knusel, K. D., Rabah, N. M., & Otteson, T. D. (2020). Health disparities in the access and cost of healthcare for otolaryngologic conditions. *Otolaryngology–Head and Neck Surgery*, 162(4), 479–488.

Seshamani, M., & Gray, A. M. (2004). A longitudinal study of the effects of age and time to death on hospital costs. Journal of health economics, 23(2), 217-235.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-experimental designs for generalized causal inference*. Wadsworth, Cengage Learning.

Showstack, J., Katz, P., Lake, J., Brown Jr, R., Dudley, R., Belle, S., Wiesner, R., Zetterman, R., Everhart, J., & Evans, R. (2000). Discount pricing and the "cost" of liver transplantation. *Liver Transplantation*, 6(1), 119–121

Kidney Transplant Cases in US: Study of Determinants of Variance in Hospital Charges

Abstract

We investigate the factors that influence the variance in hospital charges for kidney transplant cases in the US. Using the AHRQ's (Agency for Healthcare Research and Quality) HCUP's (Hospital Cot and Utilization Project) NIS (National Inpatient Sample) database, we find that variance in hospital charges is driven by patient demographics and hospital variables. We find that variance in hospital charges is determined by patient-specific factors including age, gender, race, and income, and hospital factors such as size, type, and location. Our results provide a deeper understanding of the non-clinical factors that contribute to increases in hospital charges for kidney transplant patients.

Keywords: Hospital Charges; Kidney Transplants; Hospital Characteristics; Patient demographics

1. Introduction

While spending more money on healthcare than all other countries in the world, the US has worse healthcare outcomes (Bradley et al. (2017)). US health care spending increased 4.6 percent to reach $3.6 trillion in 2018, a faster growth rate than the rate of 4.2 percent in 2017. The share of the economy devoted to health care spending declined to 17.7 percent in 2018, compared to 17.9 percent

in 2017. The main reason for this acceleration is growth in both private health insurance and Medicare, which were influenced by the reinstatement of the health insurance tax. For personal health care spending which accounted for 84 percent of national health care spending, growth in 2018 remained unchanged from 2017 at 4.1 percent. In 2018, the total number of uninsured people increased by 1.0 million to reach 30.7 million in 2018 (Hartman et al., 2020).

Costs associated with chronic kidney disease hospitalizations in the United States vary greatly among hospitalized patients (Pollock et al., 2022). Invasive procedures are a must in patients hospitalized for kidney transplants and significantly increase hospitalization costs. Moreover, the average cost of a kidney transplant in the United States is one of the most expensive single organ transplant procedures (Bentley, 2017). Based on the above, it is evident that kidney transplant hospitalizations are a major financial cost to healthcare systems that need further investigation.

2. Literature review

According to the CDC, kidney disease, also known as chronic kidney disease, is a leading cause of death in the United States. It causes more deaths than breast cancer and prostate cancer (NVS 2021 report of 2018 data). In 2021, about 37 million US adults are estimated to have CKD, and most are undiagnosed. Forty percent of people that have severely reduced function and are not on dialysis are not aware of having CKD. The end-stage renal disease (ESRD) occurs when the kidneys fail, and the patient is treated with dialysis or kidney transplant. In the United States, diabetes and high blood pressure are the two main causes of kidney disease. In 2019, treating Medicare beneficiaries with CKD cost $87.2 billion, while treating people with ESRD cost an additional $37.3 billion (CDC, 2022).

Kidney disease is a leading cause of death for people of all racial and ethnic groups in the United States, including African American, American Indian, Alaska Native, Hispanic, and white men. For women from the Pacific Islands and Asian American, American Indian, Alaska Native, and Hispanic women. It is more common in people aged 65 years or older (38%), more common in women (14%) than men (12%), and more common in non-Hispanic black adults (16%) than in non-Hispanic White adults (19%) or non-Hispanic Asian adults (13%). (Chronic Kidney Disease Facts | cdc.gov).

Previous studies show alternative determinants associated with hospitalization and hospital charges. For example, Holland et al. (2000) use demographic, clinical, and biochemical factors to predict hospitalization in a cohort of pre-dialysis patients. They find that advanced age, comorbid cardiovascular illness and anemia are independent predictors of non-elective hospitalization prior to dialysis initiation. Schrauben et al. (2020) use multivariate-adjusted Poisson regression to identify clinical and nonclinical factors associated with hospitalization rates for participants enrolled in the Chronic Renal Insufficiency Cohort (CRIC) Study. They find that adults with CKD had a higher hospitalization rate than the general population that is hospitalized. The higher hospitalization was significantly associated with clinical factors (cardiovascular disease (31.8%), genitourinary (8.7%), digestive (8.3%), endocrine, nutritional or metabolic (8.3%), and respiratory (6.7%)) and nonclinical factors (age, race/ethnicity, and gender).

Chen et al. (2017) find that the costs of EOL care for patients with CKD were driven by physician characteristics, facility factors, payment policies, and individual patient characteristics. Dai et al. (2021) use random forest and least absolute shrinkage and selection operator regression models (LASSO) to predict hospitalization expenses for inpatients with CRF. They find that hospitalization expenses are significantly related to major procedures, medical payment methods,

hospitalization frequency, length of stay, number of other diagnoses, and number of procedures. Ozieh et al. (2017) examined the trends in healthcare expenditure in adults with chronic kidney disease (CKD) and other kidney diseases (OKD) in the U.S. from 2002 to 2011. They find that race/ethnicity, hospital location, patient location, marital status, gender, education, insurance type, and income were important predictors of CKD and associated expenditures. They conclude that CKD and OKD are significant cost-drivers that impose a significant economic burden to the US population. Smith et al. (2004) find that CKD doubles the costs to the health care system and that comorbidities related to CKD contribute more to the cost of managing these patients than does CKD itself. They suggest the need to better manage the comorbid conditions to reduce medical care costs. Honeycutt et al. (2013) show that the economic burden of CKD is higher among the older adult population. The earlier stages of CKD contributed the most costs suggesting the need for early identification to better manage these costs.

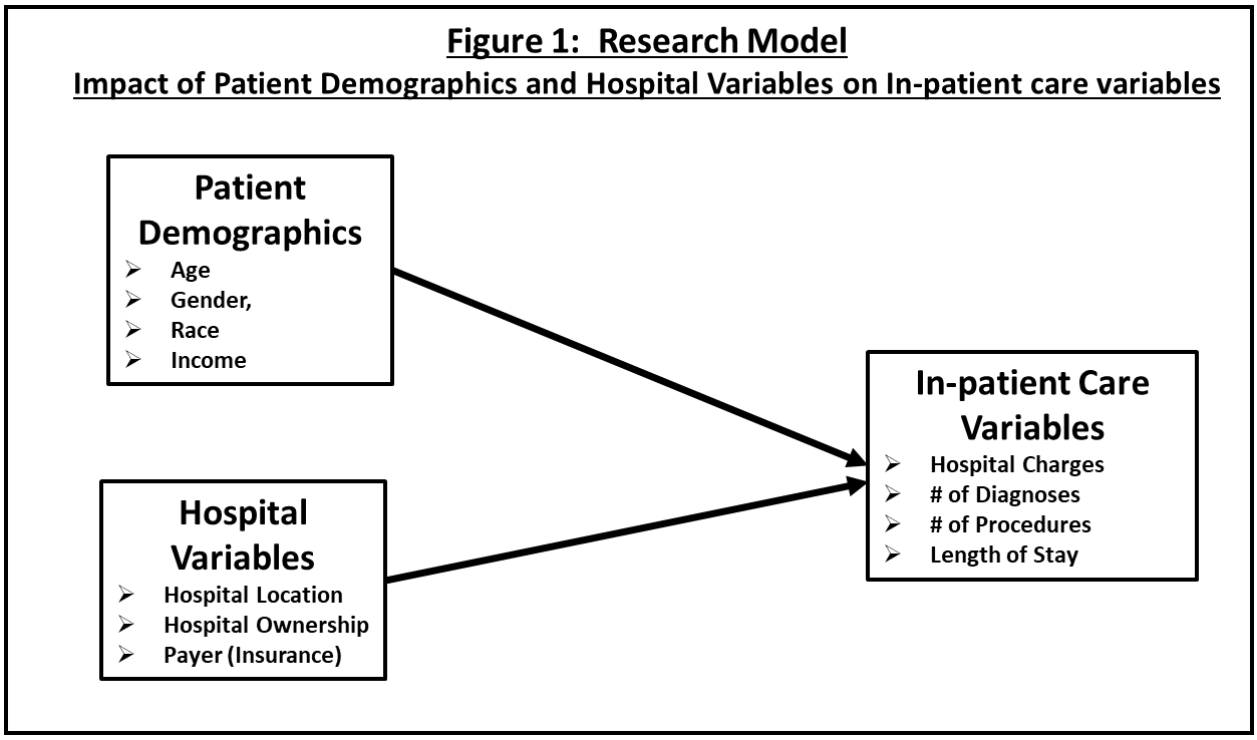
Our research makes two significant departures from many previous studies: First, we avoid linking databases from multiple sources that might introduce disparities in using data due to data collection and labeling inconsistencies. Second, we use a national stratified random sample rather than from a single source that might not represent the true national costs. Third, we limit ourselves to structural (non-clinical) rather than clinical factors influencing costs. We are not completely ignoring the clinical factors, but these are reflected in our model through structural equivalents such as the number of in-patient diagnoses and treatments.

Our research aims to explain the variance in hospital charges and in-patient kidney disease cases across the United States. AHRQ's mission is to facilitate research in healthcare costs and processes. Without going into the complexities of clinical decision-making, we wish to examine the AHRQ data to empirically reveal the impact of non-clinical variables in the AHRQ data set on

hospital charges and in-patient care. Our research focus is thus limited by the data elements in the AHRQ data set, and HIPAA requirements prohibit us from linking this data with other data sets available.

3. Grounded Research Model and Hypotheses

Our research is grounded in the empirical data available in the AHRQ data set. The AHRQ data set for 2019 in the US shows that for a sample of 4,213 kidney transplant hospitalizations, the hospital charges ranged from a minimum of $18,745 to a maximum of $1,958,373 with a mean value of $269,989. For the same sample data, the number of diagnoses ranged from 2 to 40 with an average of 14.81; the number of treatments ranged from 0 to 25 with an average of 2.79 and the length of stay in hospital ranged from 0 to 135 with an average stay of 5.88 days. Given such wide and unexplained variance, our research model examines the non-clinical variables in the AHRQ data set to explain the variance in hospital charges and in-patient care (length of stay in the hospital, # of diagnoses and # of procedures). Therefore, our research model, grounded in data available in the AHRQ data set, is depicted in Figure 1 below.



As noted earlier, we are limited to using only the patient data within the AHRQ data set. We aim to explain the variance in hospital charges and in-patient care by using the non-clinical variables in the data set. Patient age, gender, income and race are the patient specific variables that will be used in our analysis. Among these, only age is a continuous variable and the others are categorical variables. Hospital division and Hospital Ownership are also categorical variables. These are described in more detail next.

4. Research sample and variables

The Agency for Healthcare Research and Quality (AHRQ) is one of twelve agencies within the United States Department of Health and Human Services. AHRQ currently has an annual budget of more than $488.8 million (FY2022) to compile open government data for healthcare research. Since the early 1990s, AHRQ's "The Health Care Cost and Utilization Project (HCUP)" has been collecting data from 4,568 hospitals which is a representative sample of hospitals across the U.S.

The unit of analysis in the HCUP databases represents a single inpatient episode, from hospital admissions to discharge. Records from VA hospitals, hospitals on Indian Reservations, and long-term care hospitals were excluded from our study. The hospitals employ a DRG (Diagnosis Related Group) code from 000 to 999 to classify each admission. In a given year, the sample consists of more than 7 million records with information for each admission on about 250 variables.

The HIPAA Privacy Rule sets national standards for patient rights concerning health information. This rule protects individually identifiable health information by establishing conditions for its use and disclosure by covered entities. The HCUP databases conform to the definition of a limited data set. A limited data set is healthcare data in which 16 direct identifiers, specified in the Privacy Rule, have been removed. Under HIPAA, review by an institutional review board (IRB) is not required for use of limited data sets.

We focused on DRG=652 (kidney transplants) for the year 2019. Our objective was to empirically understand if non-clinical factors such as the patient demographic variables (race, income, age, and gender) impact hospitalization charges and in-patient care. We did not question the clinical decisions made by the doctors who diagnose and treat patients with diligent care providing highly patient-specific care (number of diagnoses, number of treatments, length of stay in the hospital,

etc.) aiming for the best patient outcomes. Our aim is to empirically examine the non-clinical predictors and implications of hospital charges and in-patient care.

Hospital charges do vary by disease, hence we focused only on kidney diseases with no complications and comorbidities for our study. As would be expected, even within such a seemingly homogeneous clinical category, there is significant variation in patient cases such as

age, gender, etc. that we need to consider in our study. Likewise, many hospital-specific variables such as size, location, type of hospital, etc. also cause variance in hospital charges.

4. Research Sample

It must be noted that our data for kidney transplants are from 2019 which is pre-Covid pandemic. The total number of patient records across all DRG codes in 2019 was 7,083,805 which had been collected from a stratified sample of 4,568 hospitals in the U.S. The hospital composition was 13% from the Northeast, 30% from the Midwest, 38% from the South, and 19% from the Western region. The hospitals in the sample were 20% government, non-federal hospitals; 64% private, not-for-profit hospitals, and 16% private, investor-owned hospitals. Of the total 7,083,805 records from 4,568 hospitals, the number of records in the database for kidney transplants (DRG code = 652) was 4,213 discharge records, which is the sample size for our research study. Table 1 presents the univariate statistics of our sample of 4,213. The variable descriptions are detailed at the AHRQ website: (https://www.hcup-us.ahrq.gov/db/nation/nis/nisdde.jsp).

6. Data analysis and results

The descriptive statistics of the research variables (mean, standard deviation, minimum and maximum values) are shown in Table 1.

Table 1: Descriptive Statistics					
	N	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
Hospital Charges	4200	$18,745	$1,958,373	$269,989	$18,745
# of Diagnoses	4213	2	40	14.81	6.00
# of Procedures	4213	1	25	2.79	2.18
Length of stay	4213	0	135	5.88	5.06
Age	4213	0	82	51.22	15.74
Gender	4213	0	1	0.4	0.49
Race	4123	1	6	2.00	1.21
Income Quartile	4155	1	4	2.46	1.13
Payer	4208	1	6	1.78	1.088
Hospital Division	4213	1	9	5.09	2.46
Hospital Ownership	219	1	3	1.89	0.47

Our research model in Figure 1 shows that hospital charges and in-patient care (# of Diagnoses, # Procedures and Length of Stay) are the dependent variables. The independent variables are patient age, gender, income, race and hospital division, Payer and Hospital Ownership. Thus, four regression models are used to examine the relationships between the dependent and independent variables, taken one by one on a uni-variate basis. The results of the uni-regression models are shown in Table 2. The results are discussed after Table 2.

Table 2: Uni-variate Regression Results

	Model 1	Model 2	Model 3	Model 4
	Dependent Variable in Regression Models 1-4			
Predictor Variable (Measure)*	**Hospital Charges**	**# of Diagnoses**	**# of Procedures**	**Length of Stay**
Age	F=.181, Not Sig.	F=264, Sig (.000)*	F=.632, Not Sig.	F= .013, Not Sig.
Gender	F=0.515, Not Sig.	F=0.00, Not Sig.	F=1.748, Not Sig.	F= .333, Not Sig.
Race	F=22.22, Sig (.000)*	F=2.515, Sig (.028)*	F=3.07, Sig (.009)*	F=6.37, Sig (.000)*
Income	F=0.584, Not Sig.	F=0.547, Not Sig.	F=.403, Not Sig.	F=1.568, Not Sig.
Payer	F=5.372, Sig (.000)*	F=30.34, Sig (.000)*	F=1.964, Not Sig.	F=10.0, Sig (.000)*
Hospital Division	F=66.44, Sig (.000)*	F=4.14, Sig (.000)*	F=6.17, Sig (.000)*	F=5.13, Sig (.000)*
Hospital Ownership	F=4.11, Sig (.018)*	F=1.008, Not Sig.	F=.453, Not Sig.	F=1.001, Not Sig.

*Note: Significant results related to the categorical measures have detailed explanations in the paper that show differences across categories relative to a baseline category.

When an independent variable is a categorical variable, dummy variables have to be created for use in the regression models. If a variable has “N” categories, then (n-1) dummy variables have to be used in the regression model with the excluded category as the reference to which the regression model results have to be compared. If the independent variable is continuous, then it can directly be used in the regression model. Therefore, the regression results are discussed next taking each independent variable at a time.

Age is a continuous variable with a mean value of 51.22 years, a modal value of 57 years and the median value of 54 years in the sample of 4,213 kidney transplants. The regression of hospital charges with age as the independent variable shows a non-significant positive beta coefficient (F=0.181, p=0.671). Interpretation of the slope of the regression means that the baseline hospital costs are $273,175 and with each increment of 1 year in age there would be a reduction of $62 which is not statistically significant. Thus, for a 65-year-old liver transplant patient, the predicted hospital charge would be $269,145 which is not a big difference from the mean value of $269,989 for kidney transplants. Similarly, age was not a statistically significant determinant of # of procedures (F=0.632, p=0.427) or length of stay (F=0.013, p=0.909). However, age was a statistically significant determinant of the number of diagnoses (F=264, p=.000) with a baseline of 10.1 diagnoses and with each increment of 1 year in age there would be an additional 0.1 diagnosis done. Thus, for a 65-year-old kidney transplant patient the estimated # of diagnoses would be 16.6.

Gender is a categorical variable with two categories (Male and Female). The sample of 4,213 patients is 59.9% male and 40.1% female. The regression of hospital charges with gender as the independent variable shows no statistically significant relationship (F=0.515, p= .473). Similarly, gender did not show any statistically significant impact on the number of diagnoses (F=0.0, P = 0.993); the number of procedures (F=1.748, p=.186) and the length of stay (F=0.333, p=.564).

Race is a categorical variable measured in the AHRQ data set in 6 categories. The sample of 4,213 patients is 44.6% White, 25.6% Black, 17.4% Hispanic, 6.5% Asian, 0.7% Native American, 3.0% Other and 2.1% missing. The regression of hospital charges with race as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as White in interpreting the regression results. The results in Table 2 show that race is statistically significant in its impact on hospital charges (F=22.22, p= 0.000). The baseline White

group was charged the least amount ($248,940) and all other races were charged statistically significant higher amounts of $21,521 for Blacks; $59,481 for Hispanics; $51,573 for Asians. Only Native Americans were charged $24,521 lower than Whites. The race was a statistically significant factor impacting the number of diagnoses (F=2.515, p =0.028); the # of procedures (F=3.067, p=0.009) and the length of stay (F=6.369, p=0.000). The baseline Whites had the highest number of diagnoses at 14.73, except the Blacks who had on, average, 0.6 more diagnoses relative to the Whites., while all other groups had a statistically significant lower numbers of diagnoses relative to the baseline Whites. Similarly, the baseline Whites had the lowest # of procedures except the Native Americans who had, on average, 0.2 less procedures relative to the Whites, while all other groups had a statistically higher number of procedures relative to the beeline Whites.

Income is a categorical variable measured in the AHRQ data set in four quartiles (0-25 percentile), (26-50 percentile), (51-75% percentile) and (76-100% percentile) of the average income of the ZIP code of the patient is coming from. The sample of 4,213 patients is 26.6% in the 1st quartile, 23.1% in the 2nd quartile, 25.4% in the 3rd quartile, 23.5% in the 4th quartile, and missing income data is 1.4%. The regression of hospital charges with income as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as the 1st quartile in interpreting the regression results. The results in Table 2 revealed no statistically significant differences in hospital charges across the income groups (F=.584, p=.625). Though statistically insignificant, the beta coefficients show a gradual progression upward indicating that patients with higher income were charged higher, though not statistically significant, amounts. The baseline lowest income group was charged $266,542. Relative to this first and lowest income quartile, the second income quartile was charged $1,660 higher, the third income quartile was charged $5,966 higher and the fourth income quartile was charged $7,482

higher. Income also showed no statistically significant impact on the number of diagnoses (F=.547, p=.650); on the # of procedures (F=.403, p=.751), and the length of stay (F=1.568, p=.195).

Payer is a categorical variable measured in the AHRQ data set in 6 categories. The sample of 4,213 patients is 61% Medicare, 5.6% Medicaid, 31.1% Private Insurance, 0.4% Self-pay, 0% No Charge, 1.8% Other and 0.1% missing. The regression of hospital charges with Payer as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as Medicare in interpreting the regression results. The results in Table 2 show that the Payer variable significantly impacts the Hospital Charges (F=5.371, p=0.000). For the baseline Medicare category the Hospital Charges were $276,388. Compared to this baseline (Medicare) group, only the Private Insurance group had a statistically significant lower amount of $21,022. All other payer groups did not have statistically significant differences relative to the Medicare Group.

With regards to # of diagnoses, the results in Table 2 show that the Payer variable significantly impacts the # of Diagnoses (F=30.335, p=0.000). The # of Diagnoses for the baseline Medicare category was 15.7 diagnoses. Compared to this baseline, all other groups of Payer had statistically significant lower # Diagnoses with the Medicaid 1.6 diagnoses lower, Private Insurance patients 2.4 diagnoses lower, Self-Pay 1.8 diagnoses lower, No Charge 5.7 diagnoses lower and Other 1.5 diagnoses lower.

Payer variable did not impact # of procedures (F=1.964, p=0.081). However, with regard to the length of stay, the results in Table 2 show that the Payer variable significantly impacts the length of stay (F=10.006, p=0.000). The length of stay for the baseline Medicare category was 6.1 days. Compared to this baseline, all other groups of Payer had statistically significant lower length of

stay with the Medicaid 0.85 days lower, Private Insurance patients 1 day lower, Self-Pay .35 days higher, No Charge 2.1 days lower and Other 0.1 day lower.

Hospital Division is a categorical variable measured in the AHRQ data set in 9 categories. Of the total 4,213 records for this DRG code (Kidney Transplants), and there were no missing values. The sample of 4,213 patients is 4.1% from New England (Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut), 15.3% Middle Atlantic (New York, Pennsylvania, New Jersey), 13.8% from East North Central (Wisconsin, Michigan, Illinois, Indiana, Ohio), 7.6% West North Central (Missouri, North Dakota, South Dakota, Nebraska, Kansas, Minnesota, Iowa), 21% South Atlantic (Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida), 4.8% East South Central (Kentucky, Tennessee, Mississippi, Alabama), 12.6% West South Central (Oklahoma, Texas, Arkansas, Louisiana), 7.2% Mountain (Idaho, Montana, Wyoming, Nevada, Utah, Colorado, Arizona, New Mexico) and 13.6% Pacific (Alaska, Washington, Oregon, California, Hawaii). We do not have the more granular state level data in the database.

The regression of hospital charges with Hospital Division as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as New England in interpreting the regression results. The results in Table 2 show that the Hospital Division variable significantly impacts the Hospital Charges ($F=66.44$, $p=0.000$). The Hospital Charges for the baseline New England category were $200,455. Compared to this New England baseline, all other groups showed statistically significant higher incremental charges. The highest statistically significant charge relative to the baseline (New England) was in the Pacific region at $178,236, and the lowest was in the South Atlantic at $37,857. The average charge in the South Atlantic was so close to that in New England (baseline) that it was not statistically insignificant.

Similarly, the Hospital Division variable impacted the # of Diagnoses (F=4.14, p=0.000). The baseline New England group had 15.4 diagnoses and all other groups had a lower number of diagnoses that ranged from 0 to 2. Also, the Hospital Division variable impacted # of procedures (F=6.172, p=0.000). For the baseline New England category was 3.16 procedures. Compared to this baseline New England group, all other categories of Hospital Division showed no statistically significant differences in # of procedures, except East South Central which showed a 0.81 lower number of procedures. The Hospital Division variable impacted length of stay (F=5.127, p=0.000). The baseline New England group had 6.9 days as length of stay on average, and all other groups had a lower number of days that ranged from 0.48 to 2.2 days as length of stay. Only the Mid Atlantic was not statistically significant from New England (baseline) and all other groups were statistically significantly lower relative to New England.

Hospital Ownership is a categorical variable measured in the AHRQ data set in 3 categories. Of the total 4,213 records for this DRG code (Kidney Transplants), 94.8% of the records had a missing value for this variable in the database leaving only 5.2% or 219 records for data analysis. The sample of 219 patients is 17.4% Govt-Non-federal, 76.7% Private-not_for_profit, and 5.9% Private-investor Owned. The regression of hospital charges with Hospital Ownership as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as Govt-Non-federal in interpreting the regression results. The results in Table 2 show that the Hospital Ownership variable significantly impacts the Hospital Charges (F=4.11, p=0.018). For the baseline Govt-Non-federal category, the Hospital Charges were $242,589. Compared to this baseline, the other two groups had a statistically significant higher incremental charge, namely, Private-not_for_profit with $37,700 and Private-investor_Owned

with $177,904. Clearly, investor owned hospitals had the highest hospital charges and the government non-federal hospitals charged the lowest amount.

With regard to # of diagnoses, the results in Table 2 show that Hospital Ownership did not impact the # of Diagnoses (F=1.008, p=.367). The # of Diagnoses for the baseline Govt-Non-federal hospitals was 16 diagnoses. Compared to this baseline, Private non-profit hospitals had 1.4 lower number of diagnoses and Private investor owned hospitals had a 1.2 lower number of diagnoses. Hospital Ownership did not impact # of procedures (F=.453, p=.636), and the length of stay, (F=1.101, p=.334). The # of procedures for the baseline Govt-Non-federal hospitals was 2.66. Private non-profit hospitals and Private investor-owned hospitals had a higher # of procedures of 0.41 and 0.57 respectively relative to the baseline. The length of stay for the baseline Govt-Non-federal hospitals was 5.5 days. Private non-profit hospitals and Private investor-owned hospitals had a length of stay of 1.02 days higher, and 0.4 days lower respectively relative to the baseline.

Multi-variate Regression Analysis

An earlier version of our manuscript did not include multivariate regression. One reviewer of that earlier version suggested that we also conduct multivariate regression analysis with all the independent variables thrown in as a full set of predictor variables. The rationale was that while the above uni-variate regression analyses reveal the effects of categories within the categorical variables on the dependent variable, one by one, the predictive power of these independent variables can only be assessed in the "presence" of all other explanatory variables. Following the reviewer feedback, we include the multivariate regression next, and interpret the results. variables can only be assessed in the "presence" of all other explanatory variables. Following the reviewer feedback, we include the multivariate regression next, and interpret the results.

Table 3: Multi-variate Regression Results (Significant beta coefficients are highlighted)				
	Model 1	**Model 2**	**Model 3**	**Model 4**
	Dependent Variable in the Multi-variate Regression Models 1-4			
Predictor Variable (Measure)*	**Hospital Charges**	**# of Diagnoses**	**# of Procedures**	**Length of Stay**
Intercept	**$228,123**	**13.76 Diagnoses**	**2.78 procedures**	**8.7 days**
Age	ß=-.019, p=0.804	ß=.082, p=0.307	ß=-.057, p=0.477	**ß=-.190, p=0.015 (-0.05 days)**
Baseline is Gender =1 (Male); the beta shown below is relative to the baseline				
Female [Gender =1]	**ß=-.185, p=0.10 (-$68,394)**	ß=-.050, p=0.509	ß=-.130, p=0.084	ß=-.120, p=0.100
Baseline is Race =1 (White); the betas shown below are relative to the baseline				
Race = 2 (Black)	ß=-.021, p=0.777	ß=-.124, p=0.113	ß=-.013, p=0.084	ß=-.068, p=0.364
Race = 3 (Hispanic)	ß=.111, p=0.158	ß=.060, p=0.469	ß=.102, p=0.214	ß=.101, p=0.206
Race = 4 (Asian)	ß=-.002, p=0.981	ß=.122, p=0.110	ß=.034, p=0.649	ß=-.047, p=0.518
Race = 5 (Native American)	ß=-.004, p=0.956	ß=.069, p=0.351	ß=-.003, p=0.963	ß=-.017, p=0.810
Race = 6 (Other)	ß=.092, p=0.189	ß=.048, p=0.517	ß=.032, p=0.665	ß=.019, p=0.793
Baseline is Income=1 (1st Quartile of Income); the betas below shown are relative to the baseline				
Income = 2 (2nd Quartile Income)	ß=-.053, p=0.526	ß=.086, p=0.328	ß=.063, p=0.471	ß=-.048, p=0.569
Income = 3 (3rd Quartile Income)	ß=-.042, p=0.618	ß=.073, p=0.412	ß=-.018, p=0.841	ß=-.030, p=0.726
Income = 4 (4th Quartile Income)	ß=-.123, p=0.169	ß=-.028, p=0.762	ß=.007, p=0.939	ß=-.069, p=0.447
Baseline is Payer =1 (Medicare); the betas shown below are relative to the baseline				
Payer = 2 (Medicaid)	ß=-.059, p=0.467	ß=-.047, p=0.577	ß=-.028, p=0.738	ß=.011, p=0.897
Payer = 3 (Private Insurance)	**ß=-.137, p=0.070 (-$52,079)**	ß=-.129 p=0.105	ß=-.086, p=0.278	**ß=-.164, p=0.033 (-1.7 days)**
Payer = 4 (Self-Pay)	ß=.002, p=0.983	ß=-.035, p=0.640	ß=-.055, p=0.463	ß=-.002, p=0.981
Payer = 5 (No Charge)	**No data**	**No data**	**No data**	**No data**
Payer = 6 (Other)	ß=-.006, p=0.933	ß=.005, p=0.952	ß=.105, p=0.163	ß=-.062, p=0.397

Baseline is Hospital Division =1 (New England); the betas shown below are relative to the baseline				
Hosp_Div = 2 (Middle Atlantic)	ß=.210, p=0.068	ß=.071, p=0.557	ß=.051, p=0.671	ß=.044, p=0.709
Hosp_Div = 3 (East North Central)	ß=.142, p=0.214	ß=.137, p=0.265	ß=.054, p=0.658	ß=.088, p=0.457
Hosp_Div = 4 (West North Central)	ß=.118, p=0.266	ß=.084, p=0.453	ß=-.004, p=0.973	ß=.083, p=0.438
Hosp_Div = 5 (South Atlantic)	ß=.118, p=0.326	ß=.031, p=0.809	ß=.074, p=0.554	ß=.011, p=0.930
Hosp_Div = 6 (East South Central)	ß=.037, p=0.669	ß=.022, p=0.809	ß=-.001, p=0.992	ß=-.019, p=0.828
Hosp_Div = 7 (West South Central)	ß=.197, p=0.092	ß=.149, p=0.223	**ß=.266, p=0.029 (+1.95 procedures)**	ß=.193, p=0.102
Hosp_Div = 8 (Mountain)	ß=.070, p=0.459	ß=.072, p=0.480	ß=.096, p=0.340	ß=-.023, p=0.814
Hosp_Div = 9 (Pacific)	**ß=.360, p=0.001 (+$194,138)**	ß=.135, p=0.247	ß=.017, p=0.881	ß=.057, p=0.12
Baseline is Hospital_Ownership =1 (Govt, non-federal); the betas shown below are relative to the baseline				
Hosp_Owner = 2 (Private, not for profit)	ß=.077, p=0.310	ß=-.101, p=0.218	ß=.055, p=0.503	ß=.051, p=0.520
Hosp_Owner = 3 (Private, investor owned)	**ß=.192, p=0.013 (+$167,851)**	ß=-.080, p=0.326	ß=-.045, p=0.584	ß=-.086, p=0.273
*Note: Significant results related to the categorical measures have detailed explanations in the paper that show differences across categories relative to a baseline category.				

Table 3 shows four multivariate regressions each with different dependent variables which are (1) Hospital Charges, (2) # of Diagnoses, (3) # of Procedures and (4) Length of Stay. Each of the categorical variables is represented in the predictors set by their respective dummy variables. Age is the only continuous independent variable in this full set of predictor variables.

For Model 1 in Table 3, the results show an intercept value of $228,123. Intercept value represents the predicted value when all the independent variables take on a value of zero. However, the categorical variables do not take on a value of zero and hence the intercept value cannot be interpreted in the traditionally accepted way. However, we can see in Model 1 results that females

are charged $68,394 less than males; patients with private insurance are charged $52,079 less relative to Medicare patients; patients in the Pacific region (Alaska, Washington, Oregon, California, Hawaii) are charged $194,138 more relative to patients in New England region (Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut); and finally patients in private, investor owned hospitals are charged $167,851 more relative to patients in Govt, non-federal hospitals. These specific statistically significant results may explain the intercept value ($228,123) in comparison with the overall mean value of kidney transplants hospital charges of $269,989. In summary, the predictor model allows a more nuanced understanding of factors that influence the hospital charges.

For Model 2 in Table 3, the results show an intercept value of 13.76 diagnoses. None of the predictor variables showed any statically significant impact on the dependent variable. The overall mean value of # Diagnoses was 14.81 diagnoses. We cannot explain this difference, but view both measures as quite different conceptually, with the mean as a naked measure of the average value of all measurements, whereas the intercept value is the predicted value when all the predictor variables take on a value of zero which in our case is not possible.

For Model 3 in Table 3, the results show an intercept value of 2.78 procedures. Intercept value represents the predicted value when all the independent variables take on a value of zero. However, the categorical variables do not take on a value of zero and hence the intercept value cannot be interpreted in the traditionally accepted way. However, in Model 3 results only the West South Central (Oklahoma, Texas, Arkansas, Louisiana) region had 1.95 procedures higher relative to the New England (Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut) region. The overall mean value of # Procedures was 2.79 procedures which is coincidentally very close to the intercept value. We cannot explain this coincidence, but view both measures as quite

different conceptually, with the mean as a naked measure of the average value of all measurements, whereas the intercept value is the predicted value when all the predictor variables take on a value of zero which in our case is not possible.

Finally, for Model 4 in Table 3, the results show an intercept value of 8.7 days. Intercept value represents the predicted value when all the independent variables take on a value of zero. However, the categorical variables do not take on a value of zero and hence the intercept value cannot be interpreted in the traditionally accepted way. However, we can see in Model 4 results age is a statistically significant determinant of length of stay with a coefficient of -0.05. This, for example, means that a 65-year-old kidney transplant patient would have a length of stay of only 5.5 days compared to a much younger (45 years old) kidney transplant patient who would stay a day longer. Similarly, patients in private, investor-owned hospitals have 1.75 days less stay relative to patients in Govt, non-federal hospitals. The overall mean value of length of stay was 5.88 days while the intercept value is 8.7 days in the multivariate regression. We view both measures as quite different conceptually, with the mean as a naked measure of the average value of all measurements, whereas the intercept value is the predicted value when all the predictor variables take on a value of zero which in our case is not possible.

Next, we discuss the above results and link them to existing literature with implications for addressing the widely observed variance in hospital charges and in-patient care.

7. Discussion, Implications and Limitations

Age has been studied in many earlier studies as a key determinant of hospital charges for a wide variety of hospitalizations. Past research found evidence that healthcare costs increase with patient age (Chinta et al., 2013; Farooqui & Farooqui, 2009; Peters, 2006; Jacobzone, 2003). Our findings

are consistent with past research on age and healthcare. Age is positively correlated with Hospital Charges, # of Diagnoses, # of procedures and length of stay, and significant for # of procedures. The median age is 54 years in the sample of 4,213 records analyzed. Thus, the implication for future research studies is to examine the geriatric segments of the patients for more preventive care rather than in-patient care to minimize the economic impact.

Gender related findings in our study do not reveal statistically significant differences in the length of stay, # of diagnoses, # of procedures and hospital charges. The impact of gender in healthcare has been demonstrated in past research studies (Daniel et al., 2018; Manuel, 2018). We do not find evidence of differences in healthcare measures for females and males. Thus, the implication is that chronic kidney disease equally impacts healthcare measures for both genders.

Race variable revealed a statistically significant impact on hospital charges, # of diagnoses, # of procedures and length of stay. Our findings are consistent with several other studies that found disparities in healthcare access based on race (Ruthberg et al., 2020; Bliss et al., 2015; Gulley et al., 2014). However, access to universal healthcare seems to mitigate racial disparities in access and quality of healthcare (Holtkamp, 2018).

Income variable in our study is a crude and aggregate measure based on the ZIP code of the patient and plugging that ZIP code in one of the four quartiles of national income. Hence, we do not believe that our findings are generalizable though some broad differences across income categories are suggested. Our results show no statistical differences between the quartiles. This implies that healthcare delivery is guided by standardized clinical protocols that are invariant of the income level of the patient.

Payer variable reveals interesting results. The baseline hospital charge for Medicare was $276,388. The hospital charge for the "Private Insurance" group was $21,022 lower than the baseline Medicare charge. The Payer variable of most groups was significantly lower for # of diagnoses and length of stay. One implication is that these findings raise an interesting topic for future research to be directed at examining the accounting practices of hospitals to bring to the surface the distinctions between charges and costs incurred at the procedural level.

The **Hospital Division** variable reveals some significant regional differences. The baseline hospital charge for New England was $200,455. Our results show that the Pacific (Alaska, Washington, Oregon, California, Hawaii) region is the most expensive at $178,236 higher than the baseline hospitalization charge for chronic kidney disease in the US. One implication is that these findings may provide some broad guidance for medical tourism to locate the least cost hospitals for chronic kidney disease cases.

Hospital Ownership indicates some significant differences in hospital charges. The baseline hospital charge for Govt-Non-Federal category was $242,589. Our results show that the Private-not_for_profit category had an insignificant difference of $37,700 over the baseline charge while the Private_investor_Owned category was more expensive with a significant difference of $177,904 over the baseline charge. This implies that hospital ownership plays an important role in determining healthcare costs for chronic kidney disease cases.

As with all research studies, our study suffers from several limitations. One limitation of our study uses only cross-sectional data (Shadish et al., 2002) and hence any temporal patterns cannot be inferred from our findings. The data is from 2019 which is one limitation of the study. For example, our study data comes from 2019 which is pre- Covid era. However, recent studies indicate many patients with severe COVID-19 are those with co-existing chronic conditions, including high blood

pressure and diabetes, and both diseases increase the risk of kidney disease.[1] Another limitation is that our variables come from the AHRQ's HCUP database; and other variables possibly affect hospital charges and in-patient care. Another limitation is inherent in the categorical measurement of many of our research variables which limits the analysis of variance in the dependent variables using more robust statistical techniques.

8. Conclusion

We use a sample of kidney transplant cases from actual data compiled in hospitals to investigate the factors that determine the variance in hospital charges for kidney transplant cases. We argue that there is still wide and unexplained variance in hospital charges and in-patient care despite increasing codification and standardization of clinical protocols for advanced procedures. Using a unique AHRQ-HCUP data set, we demonstrate that this variance in hospital charges and in-patient care for kidney transplant cases can result from patient-specific and hospital non-clinical factors.

Our research supports the mission of the medical profession which is to treat all patients well irrespective of their age, gender, income and race. There is a dearth of data-driven research on kidney disease hospitalizations revealing significant non-clinical determinants of hospital charges and in-patient care. Our study addresses this gap. In the early 20th century, Mary Parker Follett emphasized that leaders must have the "ability to grasp a total situation, i.e., see a whole, not a mere kaleidoscope of pieces" (Graham, 1996, p.168). Our research highlights the non-clinical variables as part of the whole picture in understanding hospital charges for kidney transplant cases.

[1] https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus/coronavirus-kidney-damage-caused-by-covid19#:~:text=Signs%20of%20kidney%20problems%20in,kidney%20injury%20may%20require%20dialysis.

References

1. Agency for Healthcare Research and Quality. Medical Expenditure Panel Survey (MEPS): household component summary tables: medical conditions, United States, Accessed April 8, 2021
2. Bliss, L. A., Yang, C. J., Kent, T. S., Ng, S. C., Critchlow, J. F., & Tseng, J. F. (2015). Appendicitis in the modern era: Universal problem and variable treatment. *Surgical Endoscopy*, 29(7), 1897–1902. https://doi.org/10.1007/s00464-014-3882-2
3. Bradley, E. H., Sipsma, H., & Taylor, L. A. (2017). American health care paradox—high spending on health care and poor health. *QJM: An International Journal of Medicine*, 110(2), 61-65. https://doi.org/10.1093/qjmed/hcw187
4. Bentley, T. S., & Hanson, S. G. (2017). 2017 US organ and tissue transplant cost estimates and discussion. *Milliman Research Report*, 1-20.
5. Brown, R. S., Lake, J. R., Ascher, N. L., Emond, J. C., & Roberts, J. P. (1998). Predictors of the cost of liver transplantation. *Liver Transplantation and Surgery*, *4*(2), 170–176.
6. Buchanan, P., Dzebisashvili, N., Lentine, K. L., Axelrod, D. A., Schnitzler, M. A., & Salvalaggio, P. R. (2009). Liver transplantation cost in the model for end-stage liver disease era: Looking beyond the transplant admission. Liver Transplantation, 15(10), 1270–1277.
7. Centers for Disease Control and Prevention, National Center for Health Statistics. About Multiple Cause 1999–2020. CDC WONDER Online Database website. Atlanta, GA: Centers for Disease Control and Prevention; 2022. Accessed February 21, 2022.
8. Chen, B., Fan, V.Y., Chou, YJ. et al. Costs of care at the end of life among elderly patients with chronic kidney disease: patterns and predictors in a nationwide cohort study. BMC Nephrol 18, 36 (2017). https://doi.org/10.1186/s12882-017-0456-2.
9. Chinta, R., Burns, D.J., Manolis, C., and Nighswander, T. (2013). Cost Creep Due to Age Creep" Phenomenon: Pattern Analyses of In-Patient Hospitalization Costs for Various Age Brackets in the United States. *Hospital Topics*, 91(4): 69-80.
10. Chinta, R., Fiedler, A., & Aeschleman, M. (2019). Variance in US Hospital Charges in Neonatal Care on the Modal DRG= 603: Data-Driven Attributions to Diagnoses, Treatments, Severity of Patient and Characteristics of Hospital. *Journal of Academy of Business and Economics,* 19(1), 5-14.
11. Dai P, Chang W, Xin Z, Cheng H, Ouyang W, Luo A. Retrospective Study on the Influencing Factors and Prediction of Hospitalization Expenses for Chronic Renal Failure in China Based on Random Forest and LASSO Regression. Front Public Health. 2021 Jun 15;9:678276. doi: 10.3389/fpubh.2021.678276. PMID: 34211956; PMCID: PMC8239170.
12. Daniel, H., Erickson, S. M., Bornstein, S. S., Kane, G. C., Gantzer, H. E., Henry, T. L., Lenchus, J. D., Li, J. M., McCandless, B. M., Nalitt, B. R., Viswanathan, L., Murphy, C. J., Azah, A. M., & Marks, L. (2018). Women's Health Policy in the United States: An American College of Physicians Position Paper. *Annals of Internal Medicine*, 168(12), 874–875. https://doi.org/10.7326/M17-3344
13. Farooqui, T., & Farooqui, A. (2009). Aging: An important factor for the pathogenesis of neurodegenerative diseases. Mechanisms of Ageing and Development, 130 (4), 203–215. https://doi.org/10.1016/j.mad.2008.11.006
14. Fountain, J. J., Chinta, R. C. & Jaramillo, H. (2020). Analysis of Hospital Charges for Heart Transplants in the US: An Empirical Comparison Across Regions and Years. JIFE, 20 (1), 73-86. dx.doi.org/10.18374/JIFE-20-1.7

15. Graham, P. Ed. (1996). Mary Parker Follett: Prophet of Management. Boston, MA: Harvard Business School Press.
16. Gulley, S. P., Rasch, E. K., & Chan, L. (2014). Difference, disparity, and disability: A comparison of health, insurance coverage, and health service use on the basis of race/ethnicity among US adults with disabilities, 2006–2008. Medical Care, 52(Supplement 3), S9–S16. https://doi.org/10.1097/MLR.0000000000000129
17. Hartman, M., Martin, A. B., Benson, J., Catlin, A., & National Health Expenditure Accounts Team. (2020). National Health Care Spending In 2018: Growth Driven By Accelerations In Medicare And Private Insurance Spending: US health care spending increased 4.6 percent to reach $3.6 trillion in 2018, a faster growth rate than that of 4.2 percent in 2017 but the same rate as in 2016. Health Affairs, 39(1), 8-17. https://.doi:10.1377/hlthaff.2019.01451
18. Holland, D. C., Lam, M. (2000). Predictors of hospitalization and death among pre-dialysis patients: a retrospective cohort study, Nephrology Dialysis Transplantation, Volume 15, Issue 5, May 2000, Pages 650–658, https://doi.org/10.1093/ndt/15.5.650
19. Holtkamp, M. D. (2018). Does race matter in universal healthcare? Stroke cost and outcomes in US military health care. Ethnicity & Health, 25(6), 888–896.
20. Honeycutt A, Segel J, Zhuo X, Hoerger T, Imai K, Williams D. Medical costs of CKD in the Medicare population. J Am Soc Nephrol. 2013;24:1478–83.
21. Jacobzone, S. (2003). Ageing and the challenges of new technologies: Can OECD social and healthcare systems provide for the future? The Geneva Papers on Risk and Insurance Issues and Practice, 28(2), 254–274. https://doi.org/10.1111/1468-0440.00222
22. Manuel, J. (2018). Racial/ethnic and gender disparities in health care use and access. *Health Services Research*, 53(3), 1407–1429. https://doi.org/10.1111/1475-6773.12705
23. National data—OPTN. (n.d.). Retrieved June 24, 2022, from https://optn.transplant.hrsa.gov/data/view-data-reports/national-data/#
24. Ozieh, M.N., Bishu, K.G., Dismuke, C.E. et al. Trends in healthcare expenditure in United States adults with chronic kidney disease: 2002–2011. BMC Health Serv Res 17, 368 (2017). https://doi.org/10.1186/s12913-017-2303-3.
25. Peters, R. (2006). Ageing and the brain. Postgraduate Medical Journal, 82(964), 84–88. https://doi.org/10.1136/pgmj.2005.036665
26. Philbin, E. F., & DiSalvo, T. G. (1998). Influence of race and gender on care process, resource use, and hospital-based outcomes in congestive heart failure. The American journal of cardiology, 82(1), 76-81. https://doi.org/10.1016/S0002-9149(98)00233-1
27. Philbin, E. F., & DiSalvo, T. G. (1999). Prediction of hospital readmission for heart failure: development of a simple risk score based on administrative data. *Journal of the American College of Cardiology*, 33(6), 1560-1566.
28. Philbin, E. F., Mccullough, P. A., Dec, G. W., & Disalvo, T. G. (2001). Length of stay and procedure utilization are the major determinants of hospital charges for heart failure. *Clinical Cardiology*, 24(1), 56-62. https://doi.org/10.1002/clc.4960240110
29. Pollock, C., James, G., Garcia Sanchez, J.J. *et al.* Cost of End-of-Life Inpatient Encounters in Patients with Chronic Kidney Disease in the United States: A Report from the DISCOVER CKD Retrospective Cohort. *Adv Ther* **39**, 1432–1445 (2022). https://doi.org/10.1007/s12325-021-02010-3
30. Reinhardt, U. E. (2006). The pricing of US hospital services: chaos behind a veil of secrecy. *Health Affairs*, 25(1), 57-69.

31. Richman, B. D., Kitzman, N., Milstein, A., & Schulman, K. A. (2017). Battling the chargemaster: a simple remedy to balance billing for unavoidable out-of-network care. *Am J Manag Care*, 23(04), e100-e105.
32. Rosenthal, E. (2014). After surgery, surprise $117,000 medical bill from doctor he didn't know. New York Times, 2014, A1.
33. Ruiz, J., Dugan, A., Davenport, D. L., & Gedaly, R. (2018). Blood transfusion is a critical determinant of resource utilization and total hospital cost in liver transplantation. *Clinical Transplantation*, *32*(2), e13164.
34. Ruthberg, J., S; Khan, H. A., Knusel, K. D., Rabah, N. M., & Otteson, T. D. (2020). Health disparities in the access and cost of healthcare for otolaryngologic conditions. *Otolaryngology–Head and Neck Surgery*, 162(4), 479–488.
35. Schrauben SJ, Chen HY, Lin E, Jepson C, Yang W, et al. (2020) Hospitalizations among adults with chronic kidney disease in the United States: A cohort study. PLOS Medicine 17(12): e1003470. https://doi.org/10.1371/journal.pmed.1003470.
36. Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-experimental designs for generalized causal inference*. Wadsworth, Cengage Learning.
37. Showstack, J., Katz, P., Lake, J., Brown Jr, R., Dudley, R., Belle, S., Wiesner, R., Zetterman, R., Everhart, J., & Evans, R. (2000). Discount pricing and the "cost" of liver transplantation. *Liver Transplantation*, 6(1), 119–121.
38. Smith DH, Gullion CM, Nichols G, et al. Cost of medical care for chronic kidney disease and comorbidity among enrollees in a large HMO population. J Am Soc Nephrol. 2004;15:1300–6.

The Non-Clinical Determinants of Variance in Hospital Charges: Evidence from Alcohol and Drug Abuse Cases in 2019 in US (DRG = 897)

Ngan Nguyen, Ph.D.
School of Business and Public Administration
University of District of Columbia
Washington, DC 20008
ngan.nguyen@udc.edu

Ravi Chinta, Ph.D.
School of Business and Public Administration
University of District of Columbia
Washington, DC 20008
Tel: (513) 262-5512
ravi.chinta@udc.edu

and

Suzan Abed, Ph.D.
School of Business and Public Administration
University of District of Columbia
Washington, DC 20008
suzan.abed@udc.edu

Keywords: Hospital Charges; Alcohol and drug abuse; Hospital Characteristics; Patient demographics; In-patient care

The Non-Clinical Determinants of Variance in Hospital Charges: Evidence from Alcohol and Drug Abuse Cases in 2019 in US (DRG = 897)

Abstract

Hospital charges for inpatient care have consistently suffered from significant variance. Efforts in healthcare reforms to make healthcare delivery efficient and effective require a deeper understanding of the factors that influence the variance in costs of healthcare delivery. While tremendous progress has been made on the clinical side of healthcare, there is still a dearth of knowledge on non-clinical factors that may contribute to variance in healthcare delivery. Our research investigates the impact of non-clinical factors, including patient characteristics and hospital characteristics, on the variance in hospital charges in alcohol and drug abuse cases in 2019 in US hospitals. We find that varying patient characteristics and hospital characteristics do affect variance in hospital charges in alcohol and drug abuse cases, contributing to the existing literature focusing on clinical factors affecting hospital charges.

Keywords: Hospital Charges; Alcohol and drug abuse; Hospital Characteristics; Patient demographics; In-patient care

Introduction

Despite various reforms for funding health care initiatives, health care costs in the U.S. have continued to increase. As per the recent National Health Expenditure Fact Sheet provided by the Centers for Medicare and Medicaid Services (CMS), the national health expenditure has grown 2.7% by the end of 2021, totaling US $4.3 trillion (approximately 18.3% of the Gross Domestic Product). On average, the U.S. spends over US $12,900 per resident per year toward health care.[1] It is considerably higher than that in other countries included in the Organization for Economic Co-operation and Development, where the average cost is only US $5,736 per person after adjusting for purchasing power.[2]

Alcohol and drug abuse addictions are extremely common problems in the United States. According to the *2020 National Survey on Drug Use and Health (NSDUH)* by the Substance Abuse and Mental Health Services Administration (SAMHSA), the COVID-19 pandemic had a negative impact on the nation's well-being. Americans responding to the NSDUH survey reported that the coronavirus outbreak adversely impacted their mental health, including by exacerbating use of alcohol or drugs among people who had used drugs in 2020. Based on data collected nationally from October to December 2020, it is estimated that 25.9 million past-year users of alcohol and 10.9 million past-year users of drugs other than alcohol reported they were using these substances "a little more or much more" than they did before the COVID-19 pandemic began. Among the 138.5 million people who were current alcohol users, 61.6 million were classified as binge drinkers and 17.7 million (28.8 percent of current binge drinkers and 12.8 percent of current alcohol users) were classified as heavy drinkers. More than 59.3 million people 12 or older used illicit drugs in 2020, including 49.6 million who used marijuana. Of the total 2.02 million people aged 12 or older using alcohol receiving treatment in 2020, 466,000 (23.1%) were treated as inpatient at a hospital. Of the total 4 million people aged 12 or older having substance use disorder receiving treatment in 2020, 801,000 (20%) were treated as inpatient at a hospital.

Hospital charges associated with alcohol use and drug abuse have been a significant contributor to rising health care costs. According to a CDC study (Sacks et al, 2015) health care expenses for treating problems caused by excessive drinking account for 11% of total cost of US $249 billion or US $28 billion in 2010. Similarly, substance abuse places a considerable burden on individuals, their families, the health care system, and the economy. The cost of treatment for substance abuse and other related illnesses is substantial. In 2003, substance abuse treatment cost the United States an estimated US $21 billion, and 77 percent of this cost was paid by public programs. In 2005, drug abuse was noted in 1.3 million hospital stays in the U.S. (about 3.3 percent of all stays), totaling US $9.9 billion in hospital costs (Kassed et al., 2007).

In a review study by Cartwright, 2008, the author examines costs as they relate to the financial costs of providing drug abuse treatment in private and public health plans, costs to society relating to drug abuse, and many smaller costing studies of various stakeholders in the health care system. The review indicates that maintaining ongoing research is highly policy relevant from the point of view of health services, and that more is needed on costing concepts and measurement applications.

[1] https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthexpenddata/nhe-fact-sheet

[2] How does health spending in the U.S. compare to other countries? - Peterson-KFF Health System Tracker

The significant and apparent variance in hospital charges for inpatient care in the U.S. has perplexed the public including many stakeholders such as the healthcare regulators and insurers. While the clinical side of inpatient care has been undergoing tremendous progress, the overall cost of healthcare has been ballooning. The purpose of this research is to conduct statistical analyses that reveal the sources of variance in hospital charges for alcohol and drug abuse cases in 2019 in terms of non-clinical factors and suggest areas for improvement for future reforms in hospital care in specific areas of patient care.

This research uses the AHRQ's (Agency for Healthcare Research and Quality) HCUP's (Hospital Cost and Utilization Project) NIS (National Inpatient Sample) database. The web site for the NIS databases is https://www.ahrq.gov/data/hcup/index.html. We use the 2019 data on alcohol and drug abuse disorders. Hospital charges are highly contingent on the nature and severity of the disease, number of diagnoses, number of treatments, length of stay and a whole series of diagnostic and treatment decisions made by well-qualified doctors or teams of doctors who make clinical decisions that are highly context- and patient- specific. Hence, it is very important to focus on one DRG code to understand the variance in hospital charges in terms of these factors. We use regression analyses on the 12,845 alcohol and drug abuse disorders in 2019 sampled in the NIS 2019 database to decompose the variance in hospital charges in terms of non-clinical factors that contribute to the charges.

Our results show that hospital charges, and other proxies including length of stay, number of diagnoses and number of treatments are correlated with patient-specific non-clinical factors (age, gender, race, and income), and the size, type, and location of the hospital. These results highlight both the success in understanding variance in hospital charges and those hospital characteristics that contribute to increases in hospital charges in alcohol and drug abuse cases.

Literature review

The United States has a serious substance misuse problem. Substance misuse is the use of alcohol or drugs in a manner, situation, amount, or frequency that could cause harm to the user or to those around them. Alcohol and drug misuse and related substance use disorders affect millions of Americans and impose enormous costs on our society. In 2015, 66.7 million people in the United States reported binge drinking in the past month and 27.1 million people were current users of illicit drugs or misused prescription drugs. The accumulated costs to the individual, the family, and the community are staggering and arise as a consequence of many direct and indirect effects, including compromised physical and mental health, increased spread of infectious disease, loss of productivity, reduced quality of life, increased crime and violence, increased motor vehicle crashes, abuse and neglect of children, and health care costs. Substance misuse problems can also result in other serious and sometimes fatal health problems and extraordinary costs; they may also lead to unexpected death from other causes. Three examples of these serious, sometimes lethal, problems related to substance misuse are driving under influence, overdose deaths, and intimate partner violence, sexual assault, and rape.

The most devastating consequences are seen in the tens of thousands of lives that are lost each year as a result of substance misuse. Alcohol misuse contributes to 88,000 deaths in the United States each year; 1 in 10 deaths among working adults are due to alcohol misuse. In addition, in 2014 there were 47,055 drug overdose deaths including 28,647 people who died from a drug overdose

involving some type of opioid, including prescription pain relievers and heroin—more than in any previous year on record.

Even though the United States spends more than any other country on health care, it ranks 27th in life expectancy, which has plateaued or decreased for some segments of the population at a time when life expectancy continues to increase in other developed countries—and the difference is largely due to substance misuse and associated physical and mental health problems. For example, recent research has shown an unprecedented increase in mortality among middle-aged White Americans between 1999 and 2014 that was largely driven by alcohol and drug misuse and suicides, although this trend was not seen within other racial and ethnic populations such as Blacks and Hispanics. An analysis from the Centers for Disease Control and Prevention (CDC) demonstrated that alcohol and drug misuse accounted for a roughly 4-month decline in life expectancy among White Americans; no other cause of death had a larger negative impact in this population.

Substance misuse and substance use disorders also have serious economic consequences, costing more than $400 billion annually in crime, health, and lost productivity. These costs are of a similar order of magnitude to those associated with other serious health problems such as diabetes, which is estimated to cost the United States $245 billion each year. Alcohol misuse and alcohol use disorders alone costs the United States approximately $249 billion in lost productivity, health care expenses, law enforcement, and other criminal justice costs. The costs associated with drug use disorders and use of illegal drugs and non-prescribed medications were estimated to be more than $193 billion in 2007. [3]

Despite decades of expense and effort focused on a criminal justice–based model for addressing substance use-related problems, substance misuse remains a national public health crisis that continues to rob the United States of its most valuable asset: its people. In fact, high annual rates of past-month illicit drug use and binge drinking among people aged 12 years and older from 2002 through 2014 emphasize the importance of implementing evidence-based public-health-focused strategies to prevent and treat alcohol and drug problems in the United States. A public health approach seeks to improve the health and safety of the population by addressing underlying social, environmental, and economic determinants of substance misuse and its consequences, to improve the health, safety, and well-being of the entire population.

Costs associated with substance misuse hospitalizations in the United States vary greatly among hospitalized patients. Peterson et al., 2021 study of assessment of annual cost of substance use disorder in US hospitals examines a total of 124,573,175 hospital emergency department encounters and 33,648,910 hospital inpatient encounters from the 2017 Healthcare Cost and Utilization Project Nationwide Emergency Department Sample and National Inpatient Sample. The research finds that total annual estimated attributable substance use disorder medical cost in hospitals was $13.2 billion and that by substance type, the cost ranged from $4 million for inhalant-related disorders to $7.6 billion for alcohol-related disorders.

Bell, 1994 investigates determinants of hospital-based substance abuse treatment programs and addresses the question of what types of hospitals are likely to engage in providing inpatient and/or outpatient treatment programs. The results indicate that organizational size (measured by the number of beds) is the best predictor of treatment service provision for both inpatient and outpatient settings, with larger hospitals being more likely to provide substance abuse programs. A need for

[3] The Surgeon General's Report on Alcohol, Drugs, and Health s prepared by the U.S. Department of Health and Human Services under the general direction of the Substance Abuse and Mental Health Services Administration https://addiction.surgeongeneral.gov/sites/default/files/surgeon-generals-report.pdf

additional chemical dependency treatment programs does not appear to be the primary motivating factor for hospitals developing this service. Rather, it seems hospitals provide these programs for other reasons--as part of providing a full array of services, as an average toward achieving organizational goals, as a means of sustaining a competitive advantage, or as a strategy for maintaining the same level of service as the competition.

Karaca and Moore, 2020 Statistical Brief shows that total costs for emergency department (ED) visits with at least one of the top five mental and substance use disorder (MSUD) diagnoses accounted for $3.9 billion in 2017, 70 percent of the $5.6 billion in total MSUD ED visit costs. In 2017, MSUD ED visits cost a total of $5,616 million. The following were the five most costly MSUD ED visit diagnoses and percentage of total MSUD ED visit costs: (1) Alcohol-related disorders ($1,220 million, 21.7 percent of the $5,616 million total MSUD ED visit costs); (2) Anxiety and fear-related disorders ($962 million aggregate costs, 17.1 percent); (3) Depressive disorders ($750 million, 13.4 percent); (4)Suicidal ideation/attempt/intentional self-harm ($612 million, 10.9 percent); (5) Schizophrenia spectrum and other psychotic disorders ($364 million, 6.5 percent). Average costs for MSUD ED visits also vary with patient characteristics and hospital characteristics.

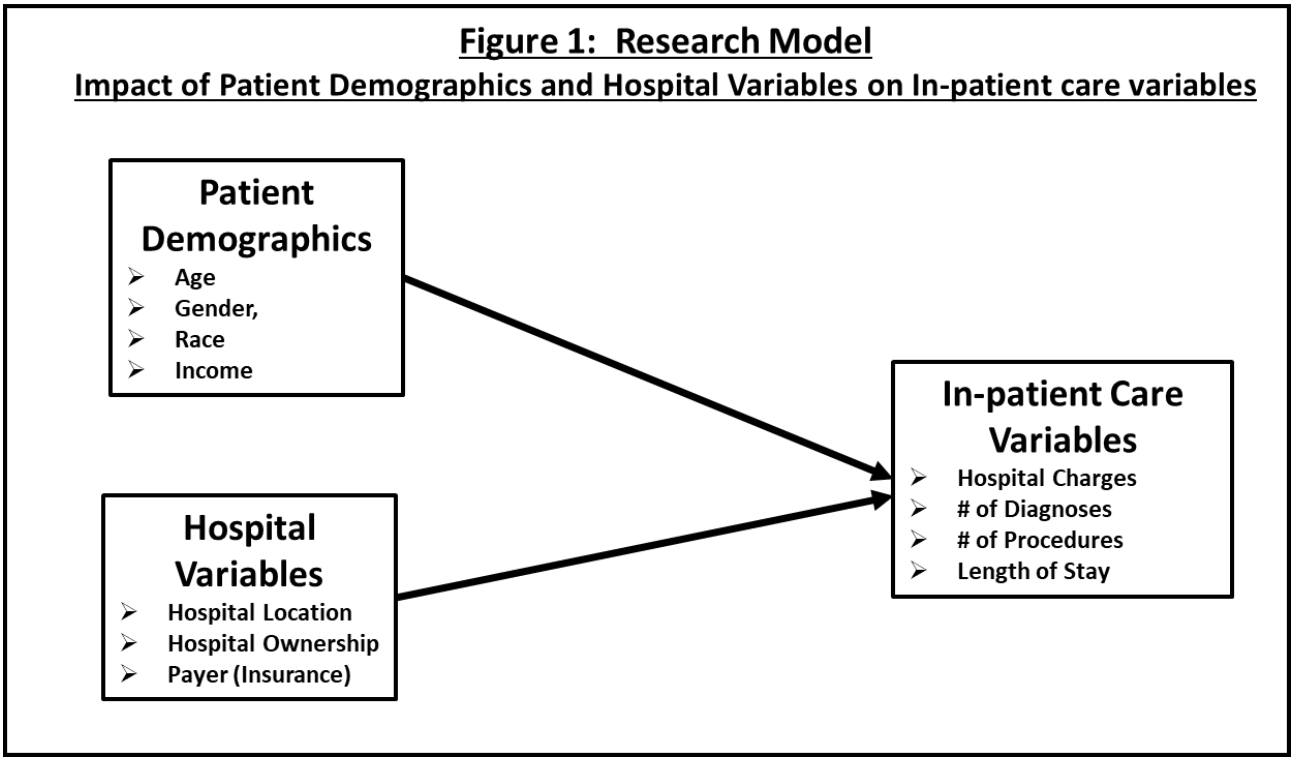
Our research makes two significant departures from many previous studies: First, we avoid linking databases from multiple sources that might introduce disparities in using data due to data collection and labeling inconsistencies. Second, we use a national stratified random sample rather than from a single source that might not represent the true national costs. Third, we limit ourselves to structural (non-clinical) rather than clinical factors influencing costs. We are not completely ignoring the clinical factors, but these are reflected in our model through structural equivalents such as the number of in-patient diagnoses and treatments.

Our research is prompted by the wide variance in hospital charges and in-patient care (length of stay in the hospital, # of diagnoses and number of procedures) that is evidenced in alcohol and drug abuse disorders across the United States. AHRQ’s mission is to facilitate research in healthcare costs and processes. Without going into the complexities of clinical decision making, we wish to examine the AHRQ data to empirically reveal the impact of non-clinical variables in AHRQ data set on the hospital charges and in-patient care (length of stay in the hospital, # of diagnoses and number of procedures). Our research focus is thus limited by the data elements in the AHRQ data set, and HIPAA requirements prohibit us from linking this data with other data sets available elsewhere.

Grounded Research Model and Hypotheses

Our research is grounded in the empirical data available in the AHRQ data set. The AHRQ data set for 2019 in the US shows that for a sample of 65,268 alcohol and drug abuse disorders hospitalizations, the hospital charges ranged from a minimum of $101 to a maximum of $804,035 with a mean value of $22,422. For the same sample data, the number of diagnoses ranged from 1 to 40 with an average of 10.25; the number of treatments ranged from 0 to 12 with an average of 0.51 and the length of stay in hospital ranged from 0 to 199 with an average stay of 4.1 days.

Given such wide and unexplained variance, our research model examines the non-clinical variables in the AHRQ data set to explain the variance in hospital charges and in-patient care (length of stay in the hospital, # of diagnoses and # of procedures). Therefore, our research model, grounded in data available in the AHRQ data set, is depicted in Figure 1 below.



As noted earlier, we are limited to using only the patient data within the AHRQ data set. Our aim is to explain the variance in hospital charges and in-patient care by using the non-clinical variables in the data set. Patient age, gender, income, and race are the patient specific variables that will be used in our analysis. Among these, only age is a continuous variable and the other three are categorical variables. Hospital division is also a categorical variable. These are described in more detail next.

Research sample and variables

The Agency for Healthcare Research and Quality (AHRQ) is one of twelve agencies within the United States Department of Health and Human Services. AHRQ currently has an annual budget of more than $488.8 million (FY2022) to compile open government data for healthcare research. Since the early 1990s, AHRQ's "The Health Care Cost and Utilization Project (HCUP)" has been collecting data from 4,568 hospitals which is a representative sample of hospitals across the U.S. The unit of analysis in the HCUP databases represents a single inpatient episode, from hospital admissions to discharge. Records from VA hospitals, hospitals on Indian Reservations, and long-term care hospitals were excluded from our study. The hospitals employ a DRG (Diagnosis Related Group) code from 000 to 999 to classify each admission. Each year, the sample consists of more than 7 million records with information for each admission on about 250 variables.

The HIPAA Privacy Rule sets national standards for patient rights with respect to health information. This rule protects individually identifiable health information by establishing conditions for its use and disclosure by covered entities. The HCUP databases conform to the definition of a limited data set. A limited data set is healthcare data in which 16 direct identifiers, specified in the Privacy Rule, have been removed. Under HIPAA, review by an institutional review board (IRB) is not required for use of limited data sets.

We focused on readmissions coded as DRG=897 (alcohol and drug abuse disorders) for the year 2019. Our objective was to empirically understand if non-clinical factors such as the patient demographic variables (race, income, age, and gender) and hospital location and payer type impact

hospitalization charges and in-patient care. We did not question the clinical decisions made by the doctors who diagnose and treat patients with diligent care providing highly patient-specific care (number of diagnoses, number of treatments, length of stay in hospital, etc.) aiming for best patient outcomes. Our aim is to empirically examine the non-clinical predictors and implications of hospital charges and in-patient care.

Different DRGs have different clinical protocols for patient care. This is the main reason why we focused only on alcohol and drug abuse disorders for our study. As would be expected, even within such a seemingly homogeneous clinical category, there is considerable variation in patient cases such as age, sex, etc. and hospital location, hospital ownership and payer type that we need to consider in our study.

Research Sample

Each record in the NIS database represents one inpatient episode. All our study variables are from the NIS 2019 database in the HCUP (https://www.hcup-us.ahrq.gov/db/nation/nis/nisdde.jsp). The total number of patient records across all DRG codes in 2019 was 7,083,805 which had been collected from a stratified sample of 4,568 hospitals in the U.S. The hospital composition was 13% from the Northeast, 30% from the Midwest, 38% from the South, and 19% from the Western region. The hospitals in the sample were 20% government, non-federal hospitals; 64% private, not-for-profit hospitals, and 16% private, investor-owned hospitals. Of the total 7,083,805 records from 4,568 hospitals, the number of records in the database for alcohol and drug abuse disorders (DRG code = 897) was 65,268 discharge records, which is the sample size for our research study. Table 1 presents the univariate statistics of our sample of 65,268. The variable descriptions are detailed at the AHRQ web site (https://www.hcup-us.ahrq.gov/db/nation/nis/nisdde.jsp).

Data analysis and results

The descriptive statistics of the research variables (mean, standard deviation, minimum and maximum values) are shown in Table 1.

Table 1: Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Hospital Charges	65,049	$101	$804,035	$22,422	$22,948
# of Diagnoses	65,268	1	40	10.25	5.2
# of Procedures	65,268	0	12	0.51	0.79
Length of stay	65,268	0	199	4.1	4.2
Age	65,262	0	90	45.20	13.700
Gender	65,263	0	1	0.30	0.460
Race	63,583	1	6	1.58	1.100
Income Quartile	62,945	1	4	2.31	1.100
Payer	65,141	1	6	2.54	1.200
Hospital Division	3,341	1	9	5.13	2.400
Hospital Ownership	3,341	1	3	2.01	0.550

The research model in Figure 1 above indicates that the hospital charges and in-patient care (length of stay in the hospital, # of diagnoses and number of procedures) are the dependent variables. The independent variables are patient age, gender, income, race and hospital division; Payer and Hospital Ownership. Thus, four regression models are used to examine the relationships between the dependent and independent variables. The results of the regression models are shown in Table 2.

Table 2: Regression Results

	Model 1	Model 2	Model 3	Model 4
	Dependent Variable in Regression Models 1-4			
Predictor Variable (Measure)*	**Hospital Charges**	**# of Diagnoses**	**# of Procedures**	**Length of Stay**
Age	F=1363, Sig (.000)*	F=6744, Sig (.000)*	F=74, Sig (0.000)*	F= 117, Sig (.000)*
Gender	F=25.6, Sig (.000)*	F=135, Sig (.000)*	F=29, Sig (0.000)*	F=1.35, Not Sig.
Race	F=92.4, Sig (.000)*	F=113, Sig (.000)*	F=43, Sig (0.000)*	F= 16.1, Sig (.000)*
Income	F=149, Sig (.000)*	F=13.7, Sig (.000)*	F=196, Sig (.000)*	F=1.1, Not Sig.
Payer	F=89.5, Sig (.000)*	F=640, Sig (.000)*	F=142, Sig (.000)*	F=99, Sig (.000)*
Hospital Division	F=40.3, Sig (.000)*	F=2.8, Sig (.004)*	F=4.6, Sig (.000)*	F=3.2, Sig (.001)*
Hospital Ownership	F=68.2, Sig (.000)*	F=4.8, Sig (.000)*	F=3.7, Sig (.024)*	F= 9.7, Sig (.000)*

*Note: Significant results related to the categorical measures have detailed explanations in the paper that show differences across categories relative to a baseline category.

When an independent variable is a categorical variable, dummy variables must be created for use in the regression models. If a variable has "N" categories, then (n-1) dummy variables must be used in the regression model with the excluded category as the reference to which the regression model results have to be compared. If the independent variable is a continuous variable, then it can directly be used in the regression model. Therefore, the regression results are discussed next taking each independent variable at a time.

Age is a continuous variable with a mean value of 45.2 years, modal value of 90 years and median value of 45 years in the sample of 65,268 alcohol and drug abuse disorders. The regression of hospital charges with age as the independent variable shows a significant negative beta coefficient (F=1363.9, p=0.000). Interpretation of the slope of the regression means that the baseline hospital costs are $11,560 and with each increment of 1 year in age there would be a reduction of $240. Thus, for a 65-year-old alcohol and drug abuse disorder patient, the predicted hospital charge would be $27,160. Similarly, age was a statistically significant (F=6744.3, p=000) determinant of the number of diagnoses with a baseline of 5 diagnoses. Age was also a statistically significant (F=74, p=.000) determinant of the number of procedures. Finally, age was a statistically significant (F=117.4, p=.000) determinant of length of stay with a baseline of 3.5 days.

Gender is a categorical variable with two categories (Male and Female). The sample of 65,268 patients is 52% male and 48% female. The regression of hospital charges with gender as the independent variable shows statistically significant impact (F=25.6, p= 0.000) on hospital charges. Similarly, gender showed statistically significant impact on the number of diagnoses (F=135.7, p =

0.000) and on the number of procedures (F= 29, p= .000). However, gender was not a statistically significant (F=1.35, p=.245) determinant of length of stay with a baseline of 4 days for the male group and no difference in length of stay for the female group.

Race is a categorical variable measured in the AHRQ data set as in 6 categories. The sample of 65,268 patients is 68.6% White,13.6% Black, 10.1% Hispanic, 0.9% Asian, 1.3% Native American, 2.9% Other and 2.6% missing data. The regression of hospital charges with race as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as White in interpreting the regression results. The results in Table 2 show that race is statistically significant in its impact on hospital charges (F=92.4, p= 0.000). The charge for the baseline White group was $22,327. Relative to the baseline White group, charges for the Blacks and Native Americans were charged $2,663 and $3,003 lower respectively, while Hispanics and Asians were charged $4,234 and $5443 higher respectively.

Race was a statistically significant factor impacting the number of diagnoses (F=113, p =0.000), but was also significant in impacting the number of procedures (F= 43.1, p=0.000). The baseline Whites had the highest number of diagnoses at 10.6, while all other groups had statistically significant lower numbers of diagnoses. The baseline Whites had the highest number of procedures at 0.5, while all other groups had statistically significant higher numbers of procedures. Race was a statistically significant factor (F=16.1, p=0.000) in the length of stay with a baseline of 4.1 days for the Whites and all other groups had statistically significant lower length of stay.

Income is a categorical variable measured in the AHRQ data set as in four quartiles (0-25 percentile), (26-50 percentile), (51-75% percentile) and (76-100% percentile) of average income of the ZIP code of the patient is coming from. The sample of 65,268 patients is 31.1% in 1st quartile, 23.2% in 2nd quartile, 23% in 3rd quartile, 19% in the 4th quartile, and missing income data is 3.6%. The regression of hospital charges with income as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as the 1st quartile in interpreting the regression results. The results in Table 2 revealed statistically significant differences across the income groups (F=149.3, p=0.000). The baseline lowest income group was charged $20,031 and relative to this baseline group, all other income groups were charged higher amounts of $1,342 for the 2nd quartile, $2,873 for the 3rd quartile and $5,123 for the 4th (highest) income quartile. Income also showed a statistically significant impact on the number of diagnoses (F=13.7, p=0.000). The baseline group of 1st quartile had the lowest number of diagnoses at 10 while the other income quartiles showed increments that gradually increased for each quartile. As for the number of procedures, there was no statistically significant difference between the first two quartiles, but the top two quartiles showed a statistically significant higher number of procedures (F=196.4, p=0.000). Finally, Table 2 results show that income did not impact length of stay (F=1.1, p=0.356).

Payer is a categorical variable measured in the AHRQ data set as in 6 categories. The sample of 65,268 patients is 16.2% Medicare, 42.2% Medicaid, 23% Private Insurance, 13.1% Self-pay, 1.5% No Charge and 4% Other. The regression of hospital charges with Payer as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as Medicare in interpreting the regression results. The results in Table 2 show that the Payer variable significantly impacts the Hospital Charges (F=89.5, p=0.000). For the baseline Medicare category, the Hospital Charges was $26,454. Compared to this baseline, all other groups except the No-Charge group had statistically significant higher incremental charges, namely, Medicaid with $4,704, Private Insurance with 5,214 and Self-Pay with $5,146. It is odd that the No-Charge group was charged a higher amount of $454 compared to the Medicare Group.

Regarding # of diagnoses, the results in Table 2 show that the Payer variable significantly impacts the # of Diagnoses (F=640, p=0.000). The # of Diagnoses for the baseline Medicare category was 12.7 diagnoses. Compared to this baseline, all other groups of Payer had statistically significant lower # Diagnoses with the Medicaid 2.6 diagnoses lower, Private Insurance patients 3.2 diagnoses lower, Self-Pay 3.5 diagnoses lower, No Charge 3.3 diagnoses lower and Other 3 diagnoses lower. Payer variable impacted # of procedures (F=142.2, p=0.000). The # Procedures for the baseline Medicare category was 0.5 procedures, and all other categories had a statistically higher number of procedures except the No-Charge group.

Regarding length of stay, the results in Table 2 show that the Payer variable significantly impacts the length of stay (F=99.4, p=0.000). The length of stay for the baseline Medicare category was 4.5 days. Compared to this baseline, Medicaid, Private-Insurance, Self-Pay and No-Charge groups had lower lengths of stay of 0.4 days, 0.66 days, 1.1 days, and 1.3 days respectively.

Hospital Division is a categorical variable measured in the AHRQ data set as in 9 categories. Of the total 65,268 records for this DRG code (alcohol and drug abuse disorders), 94.9% of the records had a missing value for this variable in the database leaving only 5.1% or 3,341 records for data analysis. The sample of 3,341 patients is 4.8% New England (Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut), 11% Middle Atlantic (New York, Pennsylvania, New Jersey), 16.6% East North Central (Wisconsin, Michigan, Illinois, Indiana, Ohio), 10% West North Central (Missouri, North Dakota, South Dakota, Nebraska, Kansas, Minnesota, Iowa), 18% South Atlantic (Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida), 5.9% East South Central (Kentucky, Tennessee, Mississippi, Alabama), 12.6% West South Central (Oklahoma, Texas, Arkansas, Louisiana), 7.8% Mountain (Idaho, Montana, Wyoming, Nevada, Utah, Colorado, Arizona, New Mexico) and 13.4% Pacific (Alaska, Washington, Oregon, California, Hawaii). We do not have the more granular state level data in the database.

The regression of hospital charges with Hospital Division as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as New England in interpreting the regression results. The results in Table 2 show that the Hospital Division variable significantly impacts the Hospital Charges (F=40.3, p=0.000). The Hospital Charges for the baseline New England category was $17,310. Compared to this New England baseline, all groups except West North Central showed statistically significant higher incremental charges. The highest was in the Mid-Atlantic region at $30,273 and the lowest was in West North Central at $15,500.

Similarly, Hospital Division variable impacted # of Diagnoses (F=2.8, p=0.004). The baseline Medicare group had 11 diagnoses and all other groups had a lower number of diagnoses that ranged from 0.05 to 1.2. Also, Hospital Division variable impacted # of procedures (F=4.6, p=0.000). The baseline New England category had 0.45 procedures. Compared to this baseline New England group, all other categories of Hospital Division showed no statistically significant differences in # of procedures, except West South Central which showed 0.21 (47% lower) number of procedures.

The Hospital Division variable also impacted length of stay (F=3.2, p= 0.001). The baseline length of stay for the New England group was 3.7 days. All other groups did not show any statistically significant differences in length of stay relative to New England.

Hospital Ownership is a categorical variable measured in the AHRQ data set as in 3 categories. Of the total 65,268 records for this DRG code (alcohol and drug abuse disorders), 94.9% of the records had a missing value for this variable in the database leaving only 5.1% or 3,341 records for data analysis. The sample of 3,341 patients is 14.7% Govt-Non-federal, 69.4% Private-not-for-profit,

15.9% Private-investor Owned. The regression of hospital charges with Hospital Ownership as the independent variable required creating dummy variables for this categorical variable and keeping the base (reference) category as Govt-Non-federal in interpreting the regression results. The results in Table 2 show that the Hospital Ownership variable significantly impacts the Hospital Charges (F=68.2, p=0.000). For the baseline Govt-Non-federal category, the Hospital Charges was $18,917. Compared to this baseline, the other two groups had a statistically significant higher incremental charge, namely, Private-not-for-profit with $4,757 and Private-investor Owned with $15,879. Clearly, investor-owned hospitals had the highest hospital charges, and the government non-federal hospitals charged the lowest amount.

Regarding # of diagnoses, the results in Table 2 show that the Hospital Ownership significantly impacts the # of Diagnoses (F=4.8, p=0.000). The # of Diagnoses for the baseline Govt-Non-federal hospitals was 9.9 diagnoses. Compared to this baseline, Private non-profit hospitals had 0.8 higher number of diagnoses and Private investor owned hospitals had 0.6 higher number of diagnoses.

For the # of procedures, the results in Table 2 show that the Hospital Ownership significantly impacts the # of procedures (F=3.7, p=0.024). The # of procedures for the baseline Govt-Non-federal hospitals was 0.26 procedures. Compared to this baseline, only Private non-profit hospitals showed a 0.09 or 34% higher number of procedures.

However, regarding length of stay, the results in Table 2 show that the Hospital Ownership significantly impacts the length of stay (F=9.7, p=0.000). The length of stay for the baseline Govt-Non-federal hospitals was 3.7 days. Compared to this baseline, only Private investor-owned hospitals showed a 0.7 or 20% lower number of days in length of stay.

Next, we discuss the above results and link them to existing literature with implications for addressing the widely observed variance in hospital charges and in-patient care.

Discussion, Implications and Limitations

Age has been studied in many earlier studies as a key determinant of hospital charges and in-patient care for a wide variety of hospitalizations. Past research found evidence attesting that healthcare costs increase with individuals' age (Chinta et al., 2013; Farooqui & Farooqui, 2009; Peters, 2006; Jacobzone, 2003). Our findings are consistent with past research on age and health. Age is positively correlated with Hospital Charges, # of Diagnoses, # of procedures and length of stay. The cost of hospitalization found in our empirical study should be highlighted as a deterrent in the messages in the advertising campaigns used in prevention programs.

Gender related findings in our study did not reveal statistically significant differences in length of stay, but gender differences were significant in hospital charges and in-patient care (# of diagnoses and # of procedures). Impact of gender in healthcare has been demonstrated in past research studies (Daniel et al., 2018; Manuel, 2018). The implication is that alcohol and drug abuse disorders impact males more than females. Future research must examine why males are more affected than females.

Race variable showed significant impact on hospital charges and in-patient care. Our findings are consistent with several other studies that found disparities in access and cost of healthcare based on race (Ruthberg et al., 2020; Bliss et al., 2015; Gulley et al., 2014). Coincidentally, research has shown that access to universal healthcare seems to mitigate racial disparities in access and quality of healthcare (Holtkamp, 2018).

Income variable in our study is a crude and aggregate measure based on the ZIP code of the patient and plugging that ZIP code in one of the four quartiles of national income. Hence, we do not believe

that our findings are generalizable though some broad differences across income categories are suggested. Our results show that only for the length of stay, there are no differences across the income groups. The myth that higher income groups are charged more is supported by our results. However, while income affects affordability of healthcare, healthcare delivery (# of diagnoses, # of procedures, length of stay and hospital charges) are guided by standardized clinical protocols that are invariant of the income level of the patient.

Payer variable reveals interesting results. The "No-Charge" is a misnomer and hospitals are in fact charging more than Medicare rates for the "No-Charge" group. Perhaps, the hospitals use this to write-off as uncovered costs. Self-Pay and Private-Insurance have the highest charges. One implication is that these findings raise an interesting topic for future research to be directed at examining the accounting practices of hospitals to bring to surface the distinctions between charges and costs incurred at the procedural level.

Hospital Division variable reveals some significant regional differences. Our results empirically identify the lowest and highest cost regions. One implication is that these findings may provide some broad guidance for medical tourism to locate least cost hospitals for alcohol and drug abuse hospitalizations.

As with all research studies, our study suffers from several limitations. One limitation of our study is the lack of an experimental design and the use of only cross-sectional data (Shadish et al., 2002). However, it is morally inadmissible to conduct experiments in high-risk procedures such as alcohol and drug abuse disorders hospitalizations. Another limitation is that we were limited by the variables available in AHRQ's HCUP database; however, there may be other variables that may affect hospital charges and in-patient care. Another limitation is inherent in the categorical measurement of many of our research variables which limits analysis of variance in the dependent variables using more robust statistical techniques.

Conclusion

We sought to build a grounded theory that is driven by actual data compiled in hospitals. While we expect that increasing codification and standardization of clinical protocols for advanced procedures would decrease the variance in hospital charges and inpatient care, there is still wide and unexplained variance. Despite the narrow data scope limited by the AHRQ-HCUP data set, we believe that our results demonstrate that variance in hospital charges and in-patient care can result from non-clinical factors. Our research supports the mission of the medical profession which is to treat all patients well irrespective of their age, gender, income and race. There is a dearth of data-driven research on specific diseases revealing statistically significant non-clinical determinants of hospital charges and in-patient care. Our study addresses this gap. In the early 20th century, Mary Parker Follett emphasized that leaders must have the "ability to grasp a total situation, i.e., see a whole, not a mere kaleidoscope of pieces" (Graham, 1996, p.168). Our research highlights the non-clinical variables as part of the whole picture in understanding hospital charges and in-patient care.

References

Bell R. Determinants of hospital-based substance abuse treatment programs. *Hosp Health Serv Adm*. 1994 Spring;39(1):93-102. PMID: 10132103.

Bliss, L. A., Yang, C. J., Kent, T. S., Ng, S. C., Critchlow, J. F., & Tseng, J. F. (2015). Appendicitis in the modern era: Universal problem and variable treatment. *Surgical Endoscopy*, 29(7), 1897–1902. https://doi.org/10.1007/s00464-014-3882-2

Brown, R. S., Lake, J. R., Ascher, N. L., Emond, J. C., & Roberts, J. P. (1998). Predictors of the cost of liver transplantation. *Liver Transplantation and Surgery*, *4*(2), 170–176.

Buchanan, P., Dzebisashvili, N., Lentine, K. L., Axelrod, D. A., Schnitzler, M. A., & Salvalaggio, P. R. (2009). Liver transplantation cost in the model for end-stage liver disease era: Looking beyond the transplant admission. *Liver Transplantation*, *15*(10), 1270–1277.

Cartwright, W. S. (2008) Economic costs of drug abuse: Financial, cost of illness, and services, *Journal of Substance Abuse Treatment*, Volume 34, Issue 2, 2008, Pages 224-233, ISSN 0740-5472, https://doi.org/10.1016/j.jsat.2007.04.003.

Chinta, R., Burns, D.J., Manolis, C., and Nighswander, T. (2013). Cost Creep Due to Age Creep" Phenomenon: Pattern Analyses of In-Patient Hospitalization Costs for Various Age Brackets in the United States. *Hospital Topics*, 91(4): 69-80.

Daniel, H., Erickson, S. M., Bornstein, S. S., Kane, G. C., Gantzer, H. E., Henry, T. L., Lenchus, J. D., Li, J. M., McCandless, B. M., Nalitt, B. R., Viswanathan, L., Murphy, C. J., Azah, A. M., & Marks, L. (2018). Women's Health Policy in the United States: An American College of Physicians Position Paper. *Annals of Internal Medicine*, 168(12), 874–875. https://doi.org/10.7326/M17-3344

Farooqui, T., & Farooqui, A. (2009). Aging: An important factor for the pathogenesis of neurodegenerative diseases. *Mechanisms of Ageing and Development*, 130 (4), 203–215. https://doi.org/10.1016/j.mad.2008.11.006

Graham, P. Ed. (1996). *Mary Parker Follett: Prophet of Management.* Boston, MA: Harvard Business School Press.

Gulley, S. P., Rasch, E. K., & Chan, L. (2014). Difference, disparity, and disability: A comparison of health, insurance coverage, and health service use on the basis of race/ethnicity among US adults with disabilities, 2006–2008. *Medical Care*, 52(Supplement 3), S9–S16. https://doi.org/10.1097/MLR.0000000000000129

Holtkamp, M. D. (2018). Does race matter in universal healthcare? Stroke cost and outcomes in US military health care. *Ethnicity & Health*, 25(6), 888–896.

Jacobzone, S. (2003). Ageing and the challenges of new technologies: Can OECD social and healthcare systems provide for the future? *The Geneva Papers on Risk and Insurance Issues and Practice*, 28(2), 254–274. https://doi.org/10.1111/1468-0440.00222

Karaca Z (AHRQ), Moore BJ (IBM Watson Health). Costs of Emergency Department Visits for Mental and Substance Use Disorders in the United States, 2017. *HCUP Statistical Brief #257*. May 2020. Agency for Healthcare Research and Quality, Rockville, MD. www.hcup-us.ahrq.gov/reports/statbriefs/sb257-ED-Costs-Mental-Substance-Use-Disorders-2017.pdf

Kassed CA, Levit KR, Hambrick MM. Hospitalizations Related to Drug Abuse, 2005. 2007 Oct. In: Healthcare Cost and Utilization Project (HCUP) Statistical Briefs [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2006 Feb-. *Statistical Brief #39*. Available from: https://www.ncbi.nlm.nih.gov/books/NBK61972/

Manuel, J. (2018). Racial/ethnic and gender disparities in health care use and access. *Health Services Research*, 53(3), 1407–1429. https://doi.org/10.1111/1475-6773.12705

National data—OPTN. (n.d.). Retrieved June 24, 2022, from https://optn.transplant.hrsa.gov/data/view-data-reports/national-data/#

National Survey on Drug Use and Health (2020) by Substance Abuse and Mental Health Services Administration. 2020 National Survey of Drug Use and Health (NSDUH) Releases (samhsa.gov)

OECD. Health at a Glance 2019: OECD Indicators. Paris, France: OECD Publishing; 2019.

Owens PL (AHRQ), Fingar KR (IBM Watson Health), McDermott KW (IBM Watson Health), Muhuri PK (AHRQ), Heslin KC (AHRQ). Inpatient Stays Involving Mental and Substance Use Disorders, 2016. *HCUP Statistical Brief #249*. March 2019. Agency for Healthcare Research and Quality, Rockville, MD. www.hcup-us.ahrq.gov/reports/statbriefs/sb249-Mental-Substance-Use-Disorder-Hospital-Stays-2016.pdf

Peters, R. (2006). Ageing and the brain. *Postgraduate Medical Journal*, 82(964), 84–88. https://doi.org/10.1136/pgmj.2005.036665

Peterson C, Li M, Xu L, Mikosz CA, Luo F. Assessment of Annual Cost of Substance Use Disorder in US Hospitals. *JAMA Netw Open*. 2021 Mar 1;4(3):e210242. doi: 10.1001/jamanetworkopen.2021.0242. PMID: 33666661; PMCID: PMC7936257.

Ruiz, J., Dugan, A., Davenport, D. L., & Gedaly, R. (2018). Blood transfusion is a critical determinant of resource utilization and total hospital cost in liver transplantation. *Clinical Transplantation*, *32*(2), e13164.

Ruthberg, J., S; Khan, H. A., Knusel, K. D., Rabah, N. M., & Otteson, T. D. (2020). Health disparities in the access and cost of healthcare for otolaryngologic conditions. *Otolaryngology–Head and Neck Surgery*, 162(4), 479–488.

Sacks JJ, Gonzales KR, Bouchery EE, Tomedi LE, Brewer RD. 2010 National and State Costs of Excessive Alcohol Consumption. *Am J Prev Med* 2015; 49(5):e73–e79.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-experimental designs for generalized causal inference*. Wadsworth, Cengage Learning.

Showstack, J., Katz, P., Lake, J., Brown Jr, R., Dudley, R., Belle, S., Wiesner, R., Zetterman, R., Everhart, J., & Evans, R. (2000). Discount pricing and the "cost" of liver transplantation. *Liver Transplantation*, 6(1), 119–121

Human-Technology Interface

IMPROVING DISPARITIES IN PROSTATE CANCER WITH TECHNOLOGY IN THE HISPANIC COMMUNITY

Crystal Wu
Pamella Howell, PhD
Eddie Castanon
California State University, Los Angeles

Abstract

In 2019, prostate cancer ranked among the top five in cancer morbidity and mortality rates in the United States. Limited health literacy and access to resources drive disparities in prostate cancer, disproportionately impacting Hispanic males. This study aims to address health inequalities by designing a low-code Mendix application prototype. We designed a prototype that optimizes patient education and improves accessibility to provider resources. The low-code mobile health application prototype functions as an equity-enhancing intervention for Hispanic men; it could yield improved outcomes for prostate cancer and reduce health disparities.

Keywords: *Prostate cancer, mHealth, Mendix, mobile applications*

Northeast Decision Sciences Institute 2023 Annual Conference

Introduction

Prostate cancer is characterized by the uncontrolled growth and spread of abnormal cells in the prostate, a small, walnut-shaped gland in males that secretes seminal fluids. Prostate cancer affects men of all ages, races, and ethnicities; however, it disproportionately affects historically marginalized groups. It is the second leading common cause of cancer deaths among Hispanic men in the United States. In 2019, Hispanic men in the United States were more likely to die from prostate cancer, with 14.3-15.6 deaths per 100,000 men (U.S. Cancer Statistics Working Group, 2022). Prostate cancer can spread aggressively but usually grows slowly. This study aims to develop a mobile application prototype that helps men locate providers and educational resources to address several social determinants that create barriers to accessing prostate cancer care.

The low code prototype in this study will facilitate early screening and detection so doctors can identify prostate cancer at a more curable stage (Carlsson & Vickers, 2020). Although no cure for prostate cancer exists, there is a way to slow its progression. By detecting prostate cancer earlier, therapy and medication can be prescribed as soon as possible. Therefore, screening and early detection can be instrumental in caring for patients with prostate cancer. The prostate-specific antigen (PSA) exam is the most widely used screening method (Owens et al., 2019).

Our mobile application prototype is an effective intervention to address the social determinants of health associated with prostate cancer disparities. Social determinants such as low health literacy, language barriers, and inadequate access to health care services create inequitable access to prostate cancer care. For example, Hispanic men lack access to prostate cancer education in their native language (Prostate Cancer Foundation, 2022). This linguistic barrier affects their health literacy and ability to access health care information. This factor and other social determinants impede access to health care, thereby affecting prostate cancer diagnosis in its early stages. We aim to alleviate the barriers underserved populations face with a prototype that improves access to providers and helps them navigate the healthcare system beginning with early screening and detection to decrease the likelihood of advanced-stage prostate cancer at diagnosis and reduce disparities in outcomes.

The mobile prototype in this study was developed using the low-code platform Mendix. The mobile application educates patients on prostate cancer and can display relevant content in

different languages, including Spanish. This study makes two significant contributions. First, we design a mobile application using the Mendix low-code platform that provides prostate cancer education and healthcare provider resources. Low-code mobile apps, like the one proposed, may enable easy deployment in provider offices. Secondly, with linguistic and cultural barriers in mind, we offer greater usability with a multilingual feature that displays content in multiple languages besides English. The resulting application addresses digital health inequities to improve health literacy, equalize healthcare access, and encourage participation in preventive services or treatment decision-making. We structure the rest of our paper as follows. The next section discusses our literature review. Then, we elaborate on our methodology and results. We include a discussion of our findings, a conclusion, and a brief discussion on future research avenues.

Literature Review

Prostate cancer-related mobile applications

More than half of US mobile phone users download health apps with their devices (Krebs & Duncan, 2015). The ubiquitous nature of mobile devices prompts people to use their phones as sources of health information. Owens et al. (2019) identified mobile prostate cancer applications and evaluated their quality, grade-level readability, cultural sensitivity, and usability heuristics to determine whether they promote informed prostate cancer screening decisions. The study found that mobile apps contained information consistent with the existing American Cancer Society's Prostate Cancer Prevention and Early Detection Guidelines but were not necessarily comprehensive enough to facilitate an informed screening decision. This gap is one catalyst that drove the development of the application in the current study.

Despite the potential to reduce prostate cancer mortality, the US Preventive Services Task Force discouraged prostate-specific antigen (PSA) screening due to concerns regarding the overdiagnosis and overtreatment of prostate cancer (Tosoian et al., 2015). Active surveillance emerged as one method to avoid overtreating indolent prostate cancer (Womble et al., 2015). Active surveillance monitors cancer progression to ensure that patients undergo medical interventions only when necessary. Wang et al. (2020) developed a mobile application for patients to assess the progression of prostate cancer risk. This application thus offers a convenient approach for patients to participate in their disease management.

Northeast Decision Sciences Institute 2023 Annual Conference

In line with the limitations of PSA testing, risk calculators were developed to determine high-grade prostate cancer risk and predict prostate cancer. De Nunzio et al. (2019) analyzed two mobile applications – the Rotterdam prostate cancer risk app and the Coral app. They validated the performance of both apps in patients undergoing prostate biopsies. They concluded that the Rotterdam app outperformed the Coral app for predicting prostate cancer or high-grade cancer diagnosis. This study represents another use case of mHealth apps in the context of prostate cancer; however, no mobile app exists to address the accessibility issues that impact health disparities among men with prostate cancer.

More than 1 in 4 cancer survivors experience a high symptom burden one year after treatment termination (Shi et al., 2011). Cancer survivors need continued symptom monitoring and management in early posttreatment survivorship. However, traditional healthcare delivery's cost and resource constraints limit the provision of services for treatment-related side effects for prostate cancer survivors (Agochukwu et al., 2018). Telemedicine-based interventions emerge to deliver survivorship care for prostate cancer survivors. Wang et al. (2021) developed a novel mHealth application called RyPros for prostate cancer survivor management. They designed this app as a patient management tool that provides reminders, assessments, and instant messaging; their study showed high acceptability among prostate cancer survivors.

Technology Use in Prostate Cancer Education

Patient education plays a critical role in managing prostate cancer. At-risk patients use mobile health technology and web resources for self-education purposes. Kobes et al. (2018) question the reliability of websites posting prostate cancer education materials due to their deficiencies in attribution, outdated content, and high readability standards. Similarly, a study by Moolla et al. (2019) determined that most websites are unreliable sources of information by themselves. Therefore, publicly available patient education content may not support informed decision-making. Patients need access to accurate, current, consistent, and objective health information. Because health apps are supplementary tools to in-person care, patients who use mobile applications and contact healthcare providers directly may make more informed medical decisions, see improved outcomes, and reduce health inequalities.

Methodology

To develop our prototype, we followed the agile methodology, an iterative application development approach (Fowler & Highsmith, 2001). This methodology promotes collaboration with users to solicit feedback as we develop our prototype. We used Mendix to develop the prototype for our mobile application. Mendix is a low-code application development platform that enables users to build and improve mobile applications at scale (Mendix, 2021). 70% of new applications organizations develop will use low-code or no-code technologies by 2025 (Wong et al., 2021). With visual modeling and minimal coding, low-code application platforms facilitate applications' rapid development and deployment. Mendix is a Model Driven Development (MDD) tool that evolved from the Computer Assisted Software Engineering (CASE) tools (Mew & Field, 2018). Therefore, Mendix users can develop fully functional applications at a higher level of abstraction from a model-driven environment, lowering the complexity of development processes.

This development has three iterations. We began the first iteration of developing our prototype in the Fall of 2021. We followed the traditional development lifecycle to create an early prototype of our application. We created our mobile application from a blank template and designed our home screen with a simple user interface. In the second iteration, we refined the entity relationship diagram (see Figure 1), a conceptual model of the database behind our application. A patient can conduct one or more searches to locate providers. They can enter the search parameters based on the provider’s name, taxonomy, language, insurance company, or location by zip code. To generate the providers listed based on each search, we visited the NPPES NPI Registry, a free online provider directory of all active National Provider Identifier (NPI) records. Individual and organizational providers receive 10-digit NPIs that uniquely identify them. We download the most recent weekly incremental NPI file and extract the zip file. To create a mobile application from the healthcare provider data file, we convert it from a comma-separated value (CSV) file to an Excel file. We then import the resulting Excel file into Mendix. We repeat this process weekly to provide users with up-to-date information that reflects newly assigned NPIs, updates to NPI data, and newly deactivated NPIs during that week. Only parts of the NPI records of healthcare providers with public relevance are published, including the provider’s name (individual or organization), specialty (taxonomy), and practice location address. Likewise, suppose the user manually enters a zip code or city or uses the location-

Northeast Decision Sciences Institute 2023 Annual Conference

sharing feature. In that case, the application will return a list of providers that satisfies the conditions in the query.

The Agency for Healthcare Research and Quality CAHPS survey (2020) shows that some patients never or only sometimes got appointments for the care they needed as soon as they needed to be seen. In our prototype, patients may choose to schedule one or more appointments. Research has shown that language barriers in healthcare lead to miscommunication between the medical professional and patient, reducing both parties' satisfaction and decreasing the quality of healthcare delivery and patient safety (Al Shamsi, 2020). Language is one of the social determinants we identified that may impact the disparities in prostate cancer. To address this barrier that Hispanic men face, our prototype allows users to choose providers who speak their language. Lastly, the American Cancer Society (n.d.) reports that many states require private health insurers to cover tests to detect prostate cancer. However, gaps in care still exist. To improve healthcare access, our search options allow patients to select providers who participate with their insurance companies.

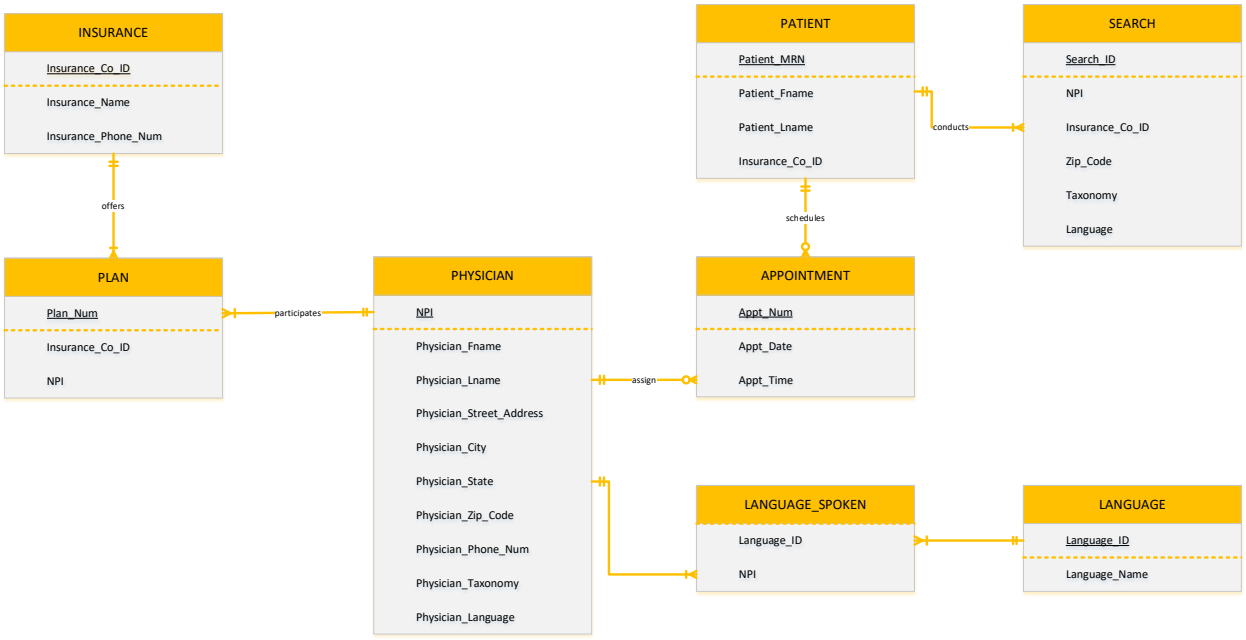


Figure 1: ER Diagram

To simulate future enhancements, we used a wireframing tool called Figma. Figma is a web application for designing user interfaces and creating prototypes for mobile devices (Putra et al., 2021). We created a wireframe for compiling education resources, improving accessibility,

and exploring our prototype's applicability for other diseases and its interoperability with electronic health records (EHR) and health information exchange (HIE). We also developed a tutorial to guide providers on using Mendix to customize our prototype to meet their patients' needs, specifically enabling patients to view our application in their preferred language.

Results

Iteration 2 Results

The prototype we developed represents a simple yet effective application for locating resources. The goal is to combat the disparity in health outcomes, specifically for Hispanic males, via universal resource accessibility. We provide screenshots of our application below. Figure 2 shows the application's home screen. Our second iteration is more patient-centered because it allows them to create personalized reminders and health-related information. From the home page, they can schedule appointments or locate providers. Patients can schedule, view, and delete appointments by navigating to the appointments page.

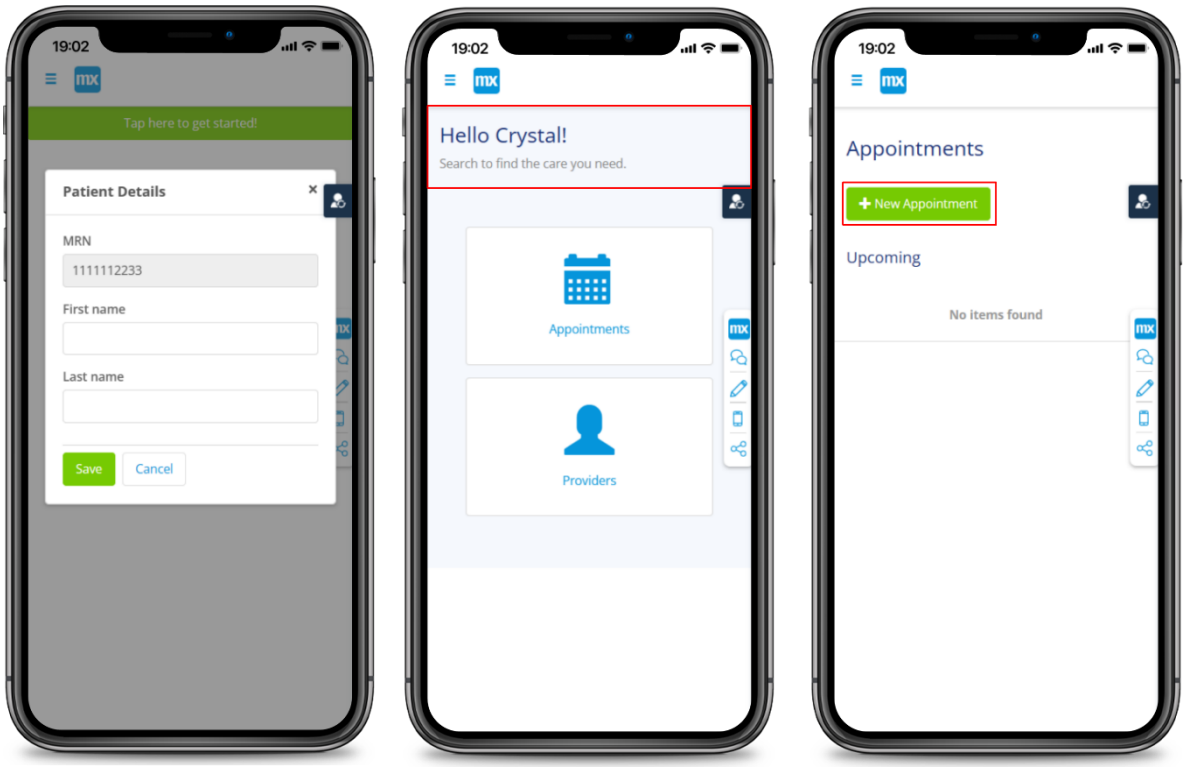


Figure 2: Home screen and Appointments page

Northeast Decision Sciences Institute 2023 Annual Conference

Patients can locate providers by entering the search parameters based on the provider's name, language, or location by zip code and receive a list of results from which they can view more specific provider details. Figure 3 illustrates a visual example of the information. The results are relatively simple, but the purpose of this application is to showcase that this low-code approach can lead to a high impact on a significant health outcome.

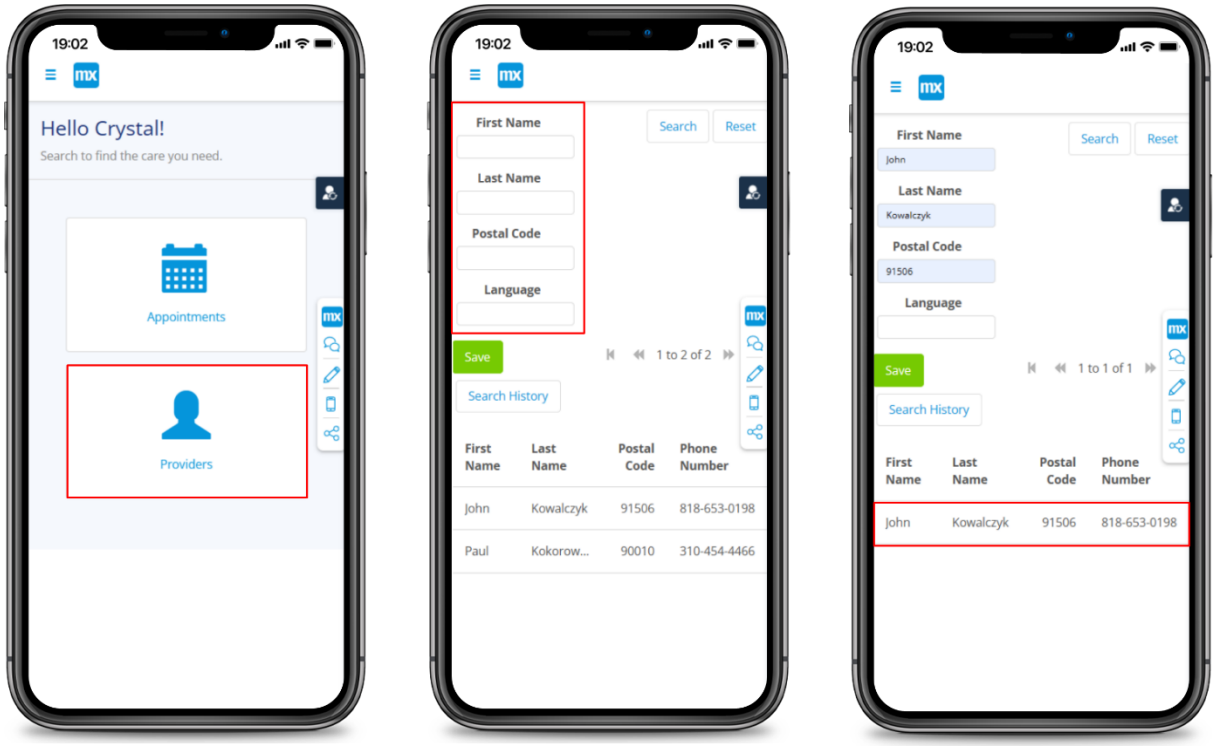


Figure 3: Provider Search and Details

Figure 4 displays the appointment scheduler. The patient will select an appropriate date and time and confirm the appointment details. They will then be redirected back to the appointments page, where they can view and edit the appointment as necessary or schedule a new one.

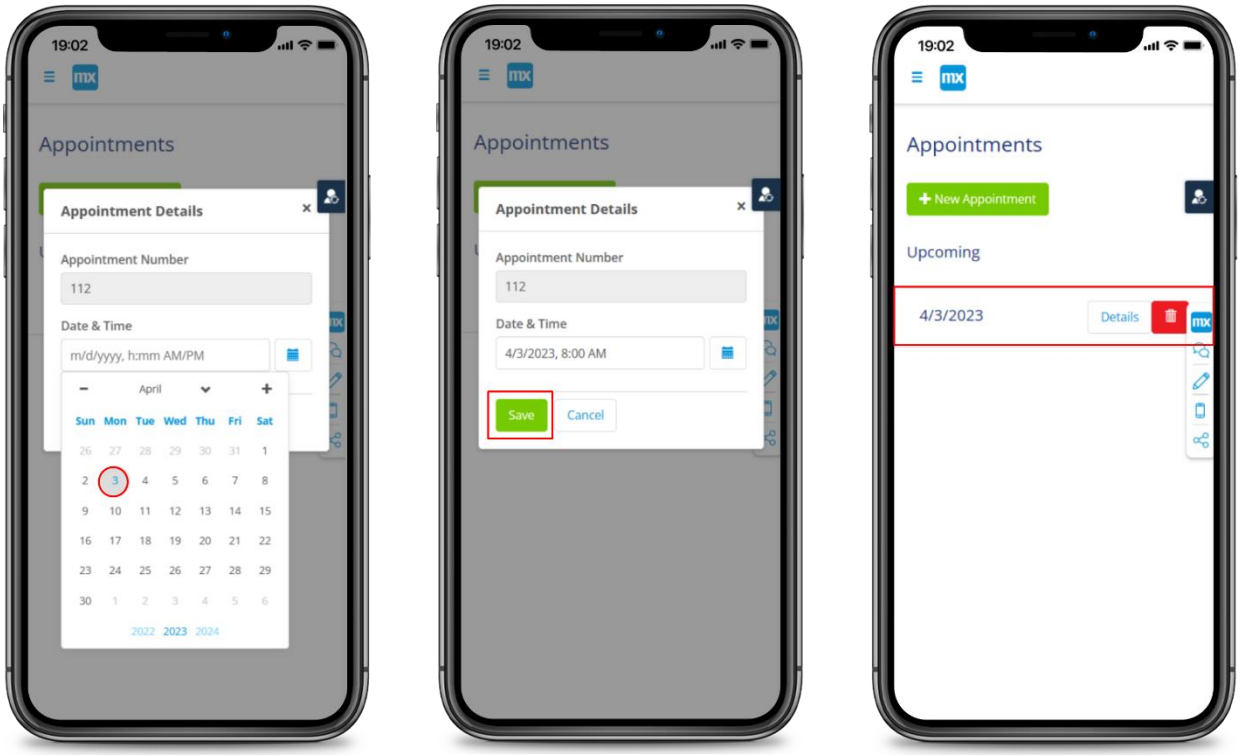


Figure 4: Appointment Scheduler

Iteration 3 Results

Figure 5(i) shows a wireframe of patient education. Patients can easily access information related to early screening and detection, diagnosis, and treatments in one convenient location, which improves their access to educational aids on prostate cancer and promotes shared decision-making on preventive care services. We will examine its interoperability with EHRs and explore the applicability of our prototype to improve disparities and health outcomes for other diseases, as illustrated in Figures 5(ii) and 5(iii).

Northeast Decision Sciences Institute 2023 Annual Conference

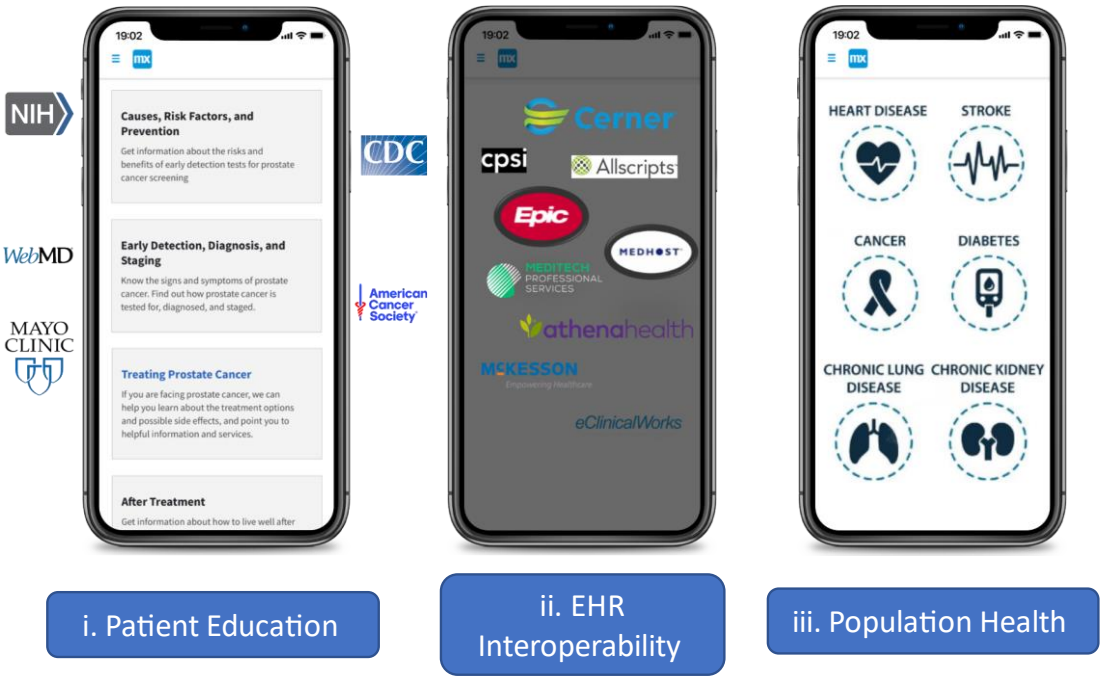


Figure 5: Wireframes for Screens in Iteration 3

Finally, Figure 6 shows a sample tutorial guiding providers on how to use the Mendix development platform to customize our prototype. To address the language barrier that Hispanic men face, we will allow patients to select their preferred language to view our application. To achieve this, we will first navigate to the language settings. We then click on the add button and add any additional languages. Mendix offers an option for each Spanish dialect. After we create a list of languages from which patients can choose, we navigate to the batch translate settings to change the language of the whole application. We will select the patients' preferred language from the destination language dropdown menu and confirm the changes.

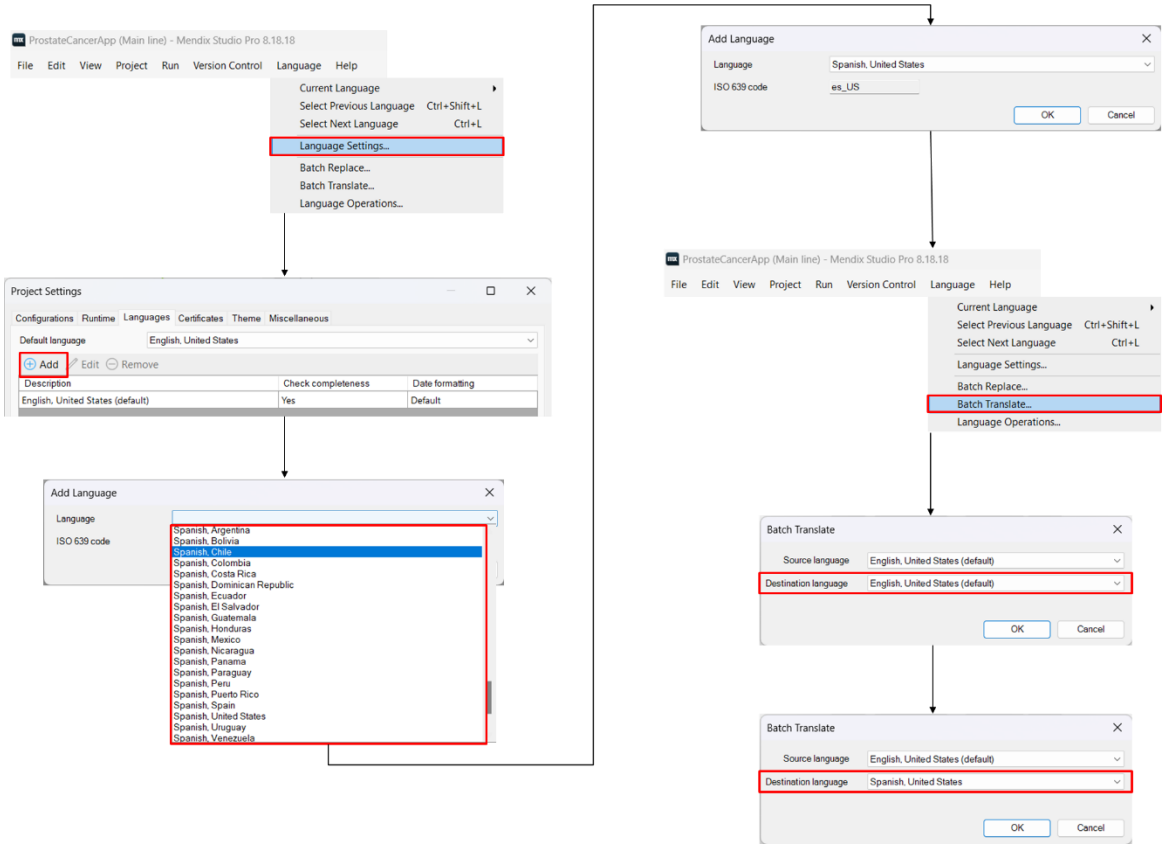


Figure 6: Sample Provider Tutorial for Language Selection

Discussion and Conclusion

We designed a low-code prototype to increase resource accessibility, especially for Hispanic men who suffer disproportionately from prostate cancer. We used Mendix, a leading application development platform solution with low-code capabilities (Wong et al., 2021). Regardless of their skill sets, users can build mobile applications native to iOS or Android operating systems in low- or no-code environments. While the conventional approach to application development requires skilled individuals, Mendix makes application development considerably more manageable, faster, and accessible for developers with varying technical expertise. Low-code development is a crucial implementation strategy, as healthcare professionals may lack the expertise to develop applications. A positive outcome of this study is that providers can modify the prototype to meet the unique needs of their facilities. Reproducibility is an inherent problem

Northeast Decision Sciences Institute 2023 Annual Conference

in research, and there are limitations to duplicating proprietary software; therefore, we can reduce disparities among patients by providing access to the prototype. Having the prototype as a resource also helps improve technology agility, which often limits technology implementation in medical facilities that cater specifically to disenfranchised individuals.

About 91% of US Hispanics reside in metropolitan areas, and although English proficiency and educational level among Hispanics have increased, educational attainment is still low, mainly among foreign-born Hispanics (Velasco-Mondragon et al., 2016). We addressed this problem by creating a user-friendly application that simplifies locating resources and allows users to specify their language preferences. The location feature assists people influenced by travel cost, time, and convenience to evaluate specific healthcare facilities; hard-to-access healthcare facilities are likely to be associated with increased morbidity and mortality (Ahmadi-Javid et al., 2017). The prototype's location function provides a solution that can benefit Hispanic males affected by prostate cancer and their community. Miscommunication in healthcare can be life-threatening. By providing access to resources to patients in their native language, we can improve the quality of healthcare and the level of satisfaction for providers and patients. Our efforts to customize the prototype to mitigate two known social determinants of health are consistent with current efforts to improve consumer health IT interventions for the US Spanish-speaking Latino population (Chaet et al., 2016).

This study focuses on developing a mobile application that helps reduce the health disparity among men diagnosed with prostate cancer. Socioeconomic factors drive the gap in prostate cancer health outcomes, with Hispanic men shouldering the burden of this disease. Prostate cancer education and resources are limited and not easily accessible for Hispanic men. Therefore, we used the Mendix low-code application development platform to design an application that identifies local resources related to prostate cancer, such as local clinics offering screening services, blood test centers, and free consultations. We developed our prototype to narrow the resources and health outcomes gap between Hispanic men and other racial or ethnic groups. For instance, the quality of care and experience would differ considerably if a Hispanic male with prostate cancer communicated with a Spanish-speaking doctor instead of one fluent in another language.

Future Research

We could enhance future versions of this app. As we develop the geolocation feature and add more functionality to this application, we will identify hotspots with a prevalence of prostate cancer diagnosis. With these hotspots located, we could recommend nearby patients with similar

health conditions such that the user can connect with them, provide or receive support, and share their personal experiences or advice with others. Nevertheless, our application and all subsequent iterations should be designed and improved to increase resource accessibility and improve the disparity in prostate cancer.

We will implement a probabilistic ranking algorithm to rank providers according to their relevance to a user's search query. We leverage mobile app analytics to enhance our search results. Our ranking algorithm calculates the provider's relevance score using data, such as frequency and recency metrics, that track visits to a location on a specific device. Higher relevance scores result in higher rankings in the search results. We will develop a recommender system to provide recommendations based on provider review site data. In either case of a location- or name-based query, the search results will provide location information with other provider attributes (provider's name, specialty, and phone number) in a configured pop-up.

In the application's related topics feature, we will source prostate cancer information from credible web resources, including the National Institutes of Health (NIH), Mayo Clinic, and WebMD. As a result, users can easily access information related to screening, diagnostics, and treatments from several objective sources in one convenient location. Users can also change the language of patient education materials; this feature allows us to deliver a linguistically tailored solution to specific groups. We also suggest additional links that direct users to external websites with more detailed information. We track the patterns of a user's navigation through the recommended sites, including the number of page views, the pattern of websites visited, the length of stay on a website, and the date and time visited. The collected data enables us to employ a hybrid recommender system that delivers user-specific content and recommends resources consumed by similar users.

Our application can potentially be impactful for the early detection of prostate cancer, especially in underserved communities. In future studies, we will use a longitudinal experimental study to collect and analyze data on the usage and benefits of the application to demonstrate efficacy. We believe the future research outlined will benefit the outcomes in the Hispanic community and males impacted by prostate cancer.

Northeast Decision Sciences Institute 2023 Annual Conference

References

Agency for Healthcare Research and Quality. (2020, March). Strategy 6a: Open access scheduling for routine and urgent appointments. https://www.ahrq.gov/cahps/quality-improvement/improvement-guide/6-strategies-for-improving/access/strategy6a-openaccess.html

Agochukwu, N. Q., Skolarus, T. A., & Wittmann, D. (2018). Telemedicine and prostate cancer survivorship: A narrative review. *Mhealth*, *4*.

Ahmadi-Javid, A., Seyedi, P., & Syam, S. S. (2017). A survey of healthcare facility location. Computers & Operations Research, 79, 223-263.

Al Shamsi, H., Almutairi, A. G., Al Mashrafi, S., & Al Kalbani, T. (2020). Implications of language barriers for healthcare: a systematic review. Oman medical journal, 35(2), e122.

American Cancer Society. (n.d.). Insurance coverage for prostate cancer screening. https://www.cancer.org/cancer/prostate-cancer/detection-diagnosis-staging/insurance-coverage.html

American Cancer Society. (2022). *Cancer Facts & Figures 2022*. cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2022/2022-cancer-facts-and-figures.pdf

American Cancer Society. (n.d.). *American Cancer Society recommendations for prostate cancer early detection*. https://www.cancer.org/cancer/prostate-cancer/detection-diagnosis-staging/acs-recommendations.html

Brewer, L. C., Fortuna, K. L., Jones, C., Walker, R., Hayes, S. N., Patten, C. A., & Cooper, L. A. (2020). Back to the future: achieving health equity through health informatics and digital health. *JMIR mHealth and uHealth*, *8*(1), e14512.

Carlsson, S. V., & Vickers, A. J. (2020). Screening for prostate cancer. *Medical Clinics*, *104*(6), 1051-1062.

Chaet, A. V., Morshedi, B., Wells, K. J., Barnes, L. E., & Valdez, R. (2016). Spanish-language consumer health information technology interventions: a systematic review. Journal of medical Internet research, 18(8), e5794

Chaudhuri, S., Das, G., Hristidis, V., & Weikum, G. (2004, August). Probabilistic ranking of database query results. In *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30* (pp. 888-899).

Coughlin, S. S. (2020). A review of social determinants of prostate cancer risk, stage, and survival. *Prostate international*, *8*(2), 49-54.

De Nunzio, C., Lombardo, R., Tema, G., Cancrini, F., Russo, G. I., Chacon, R., ... & Tubaro, A. (2019). Mobile phone apps for the prediction of prostate cancer: External validation of the Coral and Rotterdam apps. *European journal of surgical oncology*, *45*(3), 471-476.

Fowler, M., & Highsmith, J. (2001). The agile manifesto. *Software development*, *9*(8), 28-35.

Kobes, K., Harris, I. B., Regehr, G., Tekian, A., & Ingledew, P. A. (2018). Malignant websites? Analyzing the quality of prostate cancer education web resources. *Canadian Urological Association Journal*, *12*(10), 344.

Krebs, P., & Duncan, D. T. (2015). Health app use among US mobile phone owners: a national survey. *JMIR mHealth and uHealth*, *3*(4), e4924.

Mayo Clinic. (n.d.). *Prostate cancer - Symptoms and causes*. https://www.mayoclinic.org/diseases-conditions/prostate-cancer/symptoms-causes/syc-20353087

Mendix. (2021, November 19). *What is Mendix?* https://www.mendix.com/evaluation-guide/what-is-mendix/

Mew, L., & Field, D. (2018). A case study on using the mendix low code platform to support a project management course. In *Proceedings of the EDSIG Conference ISSN* (Vol. 2473, p. 3857).

Moolla, Y., Adam, A., Perera, M., & Lawrentschuk, N. (2019). 'Prostate Cancer'Information on the Internet: Fact or Fiction?. *Current urology*, *13*(4), 200-208.

Oduro, C., Connor, S. E., Litwin, M. S., & Maliski, S. L. (2013). Barriers to prostate cancer care: affordable care is not enough. *Qualitative Health Research*, *23*(3), 375-384.

Owens, O. L., Beer, J. M., Reyes, L. I., & Thomas, T. L. (2019). Systematic review of commercially available mobile phone applications for prostate cancer education. *American Journal of Men's Health*, *13*(1), 1557988318816912.

Prostate Cancer Foundation. (2022, September 20). *Prostate cancer in U.S. Hispanic men*. https://www.pcf.org/c/prostate-cancer-in-u-s-hispanic-men/

Putra, Z. F. F., Ajie, H., & Safitri, I. A. (2021). Designing a user interface and user experience from Piring Makanku application by using Figma application for teens. *IJISTECH (International Journal of Information System and Technology)*, *5*(3), 308-315.

Shi, Q., Smith, T. G., Michonski, J. D., Stein, K. D., Kaw, C., & Cleeland, C. S. (2011). Symptom burden in cancer survivors 1 year after diagnosis: a report from the American Cancer Society's Studies of Cancer Survivors. *Cancer*, *117*(12), 2779-2790.

Tosoian, J. J., Mamawala, M., Epstein, J. I., Landis, P., Wolf, S., Trock, B. J., & Carter, H. B. (2015). Intermediate and longer-term outcomes from a prospective active-surveillance program for favorable-risk prostate cancer. *Journal of Clinical Oncology*, *33*(30), 3379.

U.S. Cancer Statistics Working Group. U.S. Cancer Statistics Data Visualizations Tool, based on 2021 submission data (1999-2019): U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; www.cdc.gov/cancer/dataviz, released in June 2022.

Velasco-Mondragon, E., Jimenez, A., Palladino-Davis, A. G., Davis, D., & Escamilla-Cejudo, J. A. (2016). Hispanic health in the USA: a scoping review of the literature. Public health reviews, 37(1), 1-27

Wang, H., Chen, L., Zhou, J., Tai, S., & Liang, C. (2020). Development of Mobile Application for Dynamically Monitoring the Risk of Prostate Cancer and Clinicopathology. *Cancer Management and Research*, *12*, 12175.

Womble, P. R., Montie, J. E., Ye, Z., Linsell, S. M., Lane, B. R., & Miller, D. C. (2015). Contemporary use of initial active surveillance among men in Michigan with low-risk prostate cancer. *European urology*, *67*(1), 44-50.

Wong, J., Iijima, K., Leow, A., Jain, A., & Vincent, P. (2021, September 20). *Magic Quadrant for Enterprise Low-Code Application Platforms*. Gartner. https://www.gartner.com/doc/reprints?id=1-27IIPKYV&ct=210923&st=sb

Innovation and Creativity

Fostering Employee and Organizational Resilience in the Digital Age: Can Higher Education Help?

Guess which company came extremely close to bankruptcy in 1997? It is hard to believe, but Apple, the most valuable company on earth, was on the verge of insolvency. It is arguably the most remarkable comeback story in all of corporate history. What saved Apple? Steve Jobs understood that the company had to be innovative, change the rules, and think outside the box if the company was going to survive. Jobs was responsible for innovations that included iMacs, iPods, iPads, iTunes Music Store, and iPhones (Shontell, 2011; Weinberger and Hartmans, 2020). The Apple story should be a lesson for all companies. Solutions to problems will require creative people who think outside the discipline (what is referred to as "thinking outside the box" often means outside the discipline), can work together, and are not constrained by the models and methods of a single field. The U.S. military calls it a VUCA world with an environment that is volatile, uncertain, complex, and ambiguous (Cotton, 2019).

Digitization is altering the rules of competition, making it very easy for rival firms to come from anywhere. Hirt and Wilmott (2014, para. 1) believe that "digitization often lowers entry barriers, causing long-established boundaries between sectors to tumble." Who would have imagined that a phone would be used as a computer, camera, entertainment apparatus, educational tool, alarm clock, stopwatch, and a GPS device (remember keeping maps in your car and camera)? It is barely used as a phone. Streaming has changed how people watch movies and

television shows and listen to music; wars are being fought with cyber and drone attacks. Several non-collegiate higher education providers, including Microsoft, Google, Coursera, and the American Museum of Natural History, offer courses and certificates and compete with traditional colleges and universities. Sears, with billions of dollars in sales from its mail-order catalogs, ultimately filed for bankruptcy with a diminishing footprint presence in the retail community because it lacked the vision and agility to do what Amazon did. It is not surprising that Jeff Bezos, CEO of Amazon, is constantly trying new ideas. He believes, "Doing things at high speed, that's the best defense against the future" (Stone, 2021, para. 8). This is the secret of survival: the ability to adapt quickly once the economic environment changes.

Many of the biggest business blunders involve companies that fell into the competency trap, which means that an "organization becomes so good at one thing, it can't learn to do anything new" (Lohr & Tejada, 2018, A1). Closely related to this problem is what the military refers to as "fighting the last war," which has become shorthand to describe the stupidity of using old strategies when fighting current battles. Just as the Industrial Revolution changed how wars are fought, digitization has changed the rules of competition. Organizations that succeed will have flexible, futuristic leaders (as well as nimble employees) who understand the importance of agility and speed and know when to change course.

The amount of information is increasing exponentially, mainly because of the internet and globalization. According to Lewis (2016), knowledge doubles every 12 months; one futurist believes that knowledge doubles every 12 hours (Ray, 2020). Friedman and Lewis (2021) posit that organizational agility is essential for survival during the information age because firms can become rapidly outdated. Companies often have to reinvent themselves completely to survive. The list of companies now in trouble or bankrupt because they were slow to adapt to changing

realities continues to grow. It includes A&P, AOL, Blackberry, Blockbuster, Borders, Compaq, Kodak, MySpace, Netscape, Polaroid, Radio Shack, Sears, Tower Records, Toys R Us, Wang Laboratories, Xerox, and Yahoo.

Some Examples of Firms that Missed Out on Huge Opportunities

Examples of firms that missed out on enormous, profitable opportunities include Xerox, Kodak, Excite, and IBM.

In the early 1970s, Xerox developed the graphical user interface (G.U.I.) that uses icons and a mouse at its PARC research and development lab. This G.U.I. allowed users to work with icons on a computer rather than text. Today, this G.U.I. is used by almost every personal computer. Unfortunately, Xerox was so focused on profits from copying machines that it did not recognize the value of its innovation. Xerox even developed a personal computer -- which they never marketed -- called the Xerox Alto that used the interface. Xerox executives were used to productivity measures like "pages copied" and did not understand how important the G.U.I. desktop would become. Against their own design team's advice, they allowed Steve Jobs and several Apple employees to visit the PARC facilities in exchange for some Apple stock. That visit enabled Apple to become a trillion-dollar company. Xerox later tried suing Apple for using the G.U.I. technology in the Apple Macintosh personal computer, but the case was dismissed (Lindner, 2008; Nazar, 2013). Jobs remarked: "If Xerox had known what it had and had taken

advantage of its real opportunities, it could have been as big as I.B.M. plus Microsoft plus Xerox combined—and the largest high-technology company in the world," (Gladwell, 2011, para. 6). Xerox, the once-powerful technology company, is the "poster child for monopoly technology businesses that cannot make the transition to a new generation of technology."

In 1999, Excite was the second largest search engine, and Google was tiny. Excite was given the opportunity to purchase Google for $750,000. Excite declined the deal because the metric that mattered to them was "stickiness", which is the ability to get users to stay on the website for more extended periods of time. The executives at Excite felt Google's search engine (back then, it was called BackRub) was too quick, which meant users would leave the website and not linger there. Excite lost an opportunity to own Google, a trillion-dollar company (Jurberg, 2021).

Kodak may also have fallen into the competency trap. It was a world brand recognized everywhere and made an enormous profit from film. The company invented the digital camera in 1977 but was reluctant to switch to newer technology because of the significant returns from its film business. By the time the company switched to digital technology, it was too late: other firms had taken over the digital market (Roesler, 2015).

In 1980, I.B.M. approached Bill Gates to develop an operating system for its personal computers. Microsoft was a small firm with 40 employees and annual revenues of $7.5 million; I.B.M. had more than 340,000 employees. Gates took the job and provided I.B.M. with the system known as PC-DOS in exchange for $80,000. There was a stipulation that the copyright for the operating system would remain with Microsoft. This allowed Microsoft to sell its operating system to other personal computer manufacturers as MS-DOS. I.B.M. did not realize

that the real money to be made was not in the hardware but in the software (Business Insurance Staff, 2011). Today, Microsoft is worth more than IBM. and has significantly greater revenues.

How to Develop Organizational Agility

Complacency is a disaster for any organization. Brosseau et al. (2019) emphasize how difficult it is to move to an agile operation model, but it must be done. Dinosaurs did not disappear because they were weak but because they could not adapt.

The first step in developing organizational agility is to eliminate silos. A silo mentality is the adversary of organizational agility, as is too much bureaucracy. A silo mentality occurs when groups or different departments in an organization refuse to share knowledge. Capable corporate leaders engage in silo-busting because silos breed tribalism (Tett, 2015). Indeed, overly large organizations often find themselves with rigid silos and are consumed with turf battles. This is not the way to create organizational agility. People from different departments must collaborate and share information to succeed in today's economy.

Step two is to become a learning organization; Peter M. Senge popularized this term in his influential 1990 book *The Fifth Discipline* (Senge 1990). Organizations, such as Amazon and Google, that want to be innovative, inventive, and nimble are becoming learning organizations (Argyris and Schoen 1996; Senge 1990.

Friedman *et al.* (2005) summarize what one should expect to see in a learning organization. First and foremost is the belief in continuous and collective learning, collaboration, knowledge sharing, and a commitment to lifelong learning. The organization must be nimble and develop the ability to adapt to changing conditions, *i.e*., an ability to regenerate, renew, and revitalize itself. It cannot allow stagnation. Also, there must be a concern for people and respect (and empowerment) for employees. The individuals who make up the organization must learn from experience and mistakes — experience is the best teacher — and learn from the experiences of others in the organization. There must be a willingness to experiment and take chances; this means that there must be a tolerance for failure. Diversity is seen as a plus since it allows for new ideas.

Step three is to hire employees that understand the importance of employee creativity and adaptability (Friedman, 2020; Friedman and Lewis, 2021). Employees must possess the ability and willingness to adapt to changing business conditions. This means that employees must know how to acquire knowledge. There was a time when the skills students gained in college had a "shelf life" of many years. This is no longer the case; to succeed in the workplace, one must know how to upgrade one's skills on a frequent and consistent basis.

Given the importance of organizational resilience, there should be an emphasis on transversal skills (also known as transferable skills). These are non-specific competencies unrelated to a particular position and can be applied to various settings and jobs. Some examples include critical thinking, information literacy, problem-solving, creativity, interpersonal skills (e.g., communication skills and ability to work in teams), intrapersonal skills (e.g., persistence), global citizenship (respect for diversity and tolerance for others), and quantitative reasoning

(Care and Luo, 2016). These skills are valued by employers and may be taught in liberal arts and business classes.

Step four is to flatten the organizational hierarchy. Organizations that want to create resilience and agility must reduce layers of bureaucracy to promote open communications, make organizations nimbler and more effective, and enhance employee involvement. A flat organizational structure is characterized by short chains of command and a wide span of control, which refers to the number of subordinates controlled by the supervisor. A tall or hierarchical organizational structure is marked by long chains of command and narrow spans of control. Each additional layer of bureaucracy slows down decision-making, increases costs, hurts creativity and leads to corporate sluggishness.

In the knowledge economy, no efficient business can afford a complicated organizational hierarchy. The hierarchical organizational structure does a superb job of maintaining the status quo and makes change difficult, which is why eliminating it is so challenging (Morgan, 2015).

What Colleges Need to Emphasize

Considering how vital organizational and employee resilience has become, colleges should rethink what needs to be stressed if the goal is to prepare graduates for the workforce.

Collaboration and Teamwork: In today's global economy, there are few jobs where people work alone. Therefore, people must learn how to work in teams, collaborate, listen to what others have to say, and share knowledge. Lee (2018) contends that the way to come up with faster solutions to problems is to bring people together from diverse backgrounds. Many of humankind's most significant accomplishments (e.g., the International Space Station) resulted from scientists from many different countries working together. It should be noted that

"Researchers who work in more than one country have 40% higher average citation rates, while internationally collaborative papers have a greater impact than domestic papers" (para. 3).

As noted above, silos can be a disaster for companies that need to innovate. Unfortunately, academic departments often encourage the ludicrous belief that one discipline is superior to others and can provide all the answers. This belief interferes with critical thinking and is a cognitive bias known as *déformation professionelle* — the tendency to see things narrowly, i.e., from the point of view of one's discipline or profession (Friedman, 2017). The turf battles in academia have become legendary (e.g., economics with sociology or liberal arts vs. business). It is doubtful that students will learn to respect and work with individuals from other disciplines, from professors whose loyalty is to their department, not the institution.

Edwards (1999) asserts, "in so many cases, the most provocative and interesting work is done at the intersections where disciplines meet, or by collaborators blending several seemingly disparate disciplines to attack real problems afresh." Klein (1996, p. 191) also observes: "Almost all significant growth in research in recent decades ... has occurred at the 'interdisciplinary borderlands' between established fields." It is unrealistic to believe that an individual from only a single discipline can provide the necessary knowledge to solve problems that arise ten or even five years later. The reality is that collaboration among several authors, especially those from different disciplines, produces better and more genuinely impactful scholarly papers than research conducted by single authors (Council of State Governments, 2015). This is why educators must teach students to respect all knowledge and understand that one must join forces with others to succeed.

To be successful in the workplace, one must possess active listening skills; a talent rarely taught in higher education. This is the ability to pay attention to what others say. It indicates

respect for what others think and is needed for effective communication. Unfortunately, few possess this skill and are likely to be distracted with their smartphone, look around, or fidget with some object rather than focus on what someone else says (Schwantes, 2022).

In a global economy, many jobs require interacting with people from countries all over the world. Numerous companies are working on building secure and resilient supply chains. In fact, because of the problems with China, some companies are moving their operations to countries such as Vietnam and India. It is not surprising that researchers are discovering that the foremost predictor of success when interacting with people from different countries or cultures is cultural intelligence (CQ). It is more important than IQ or EQ (emotional intelligence). To communicate properly, people must appropriately use their body language, gestures, facial expressions, tone of voice, posture, and mannerisms (Robson, 2017).

Critical Thinking: According to Dyer (2011, p. 2), "critical thinking is an approach to reading, thinking, and learning that involves asking questions, examining our assumptions, and weighing the validity of arguments." Critical thinkers are "self-aware, curious, and independent. They introspect on their own thinking processes; they work at knowing their own biases and can name the strategies they are using when they make judgments (self-aware)." It can be taught in courses ranging from literature, history, philosophy to mathematics, accounting, and statistics, etc. There is no one discipline that "owns" critical thinking (Ennis, 1997; Friedman, Frankel, & Friedman, 2016). It can be taught to enhance problem-solving and reasoning skills, regardless of major. Given how technology is becoming more complex and sophisticated, employees need to have the ability to create innovative solutions to complex problems. Admittedly, some believe that it is not clear that critical thinking is a skill that can be taught (Schlueter, 2016). In any case, it should probably not be a separate course but part of every course. Academics talk about ethics

across the curriculum; the same should be done for critical thinking, and it should be spread over the entire curriculum.

The push towards diversity in education rarely includes the diversity of ideas and thus runs counter to critical thinking. Courses that should encourage independent thought often serve to indoctrinate students so as to conform to the instructor's opinions and ideas. Such classes may actually be harmful. Critical thinking is not about what to think but how to think. Even a professor who is an avowed Marxist should be presenting both sides of an argument. The goal should be to teach students to consider all sides of an argument and then make a rational decision. Indoctrination courses are popular in totalitarian regimes but have no place in a democracy. A recent study found that 63.5 % of students believe "the climate on my campus prevents some people from saying things they believe because others might find them offensive" (Burt, 2022, para. 3). Paul (2022) posits that there is a mistaken belief that minorities who are elected have a special responsibility to speak for their own ethnic group. It is becoming evident that diversity is "more than skin deep," and there is a great deal of variety of political opinion among minorities. It is not impossible to have a very conservative Supreme Court in the United States consisting of Blacks and Latinos.

Freedom of speech is being suppressed or limited on many campuses because the cancel culture has become widespread. Even professors are afraid of expressing opinions that are not seen as "woke" because they might be fired. Thanks to social media, it has become relatively easy to find out if someone articulates a viewpoint that might be inconsistent with the "woke" establishment.

Creative Thinking: Creative thinking engages the right side of the brain and is not the same as critical thinking. Creativity is not only crucial in art and music; it is a vital skill, and

employers want to hire creative thinkers. One survey of 1,500 executives found that they believed that creativity was the most essential business skill in the modern world (cited in Davis, 2018). Teaching creativity is being stressed at many K-12 schools (Davis, 2018). Friedman, Friedman, and Leverton (2016) provide compelling evidence that diversity is essential for augmenting creativity, i.e., having people from various backgrounds with diverse opinions work together on projects. In a knowledge economy, creativity requires that people possess the ability to collaborate. Great ideas arise in teams--think of the Manhattan Project.

Solutions to problems will require creative people who think outside the discipline (what is referred to as "thinking outside the box" often means outside the discipline), have the ability to work together productively, and are not constrained by the models and methods of a single field.

Passion for Learning: One of the buzzwords one hears in academia and the corporate world is lifelong learning. If higher education institutions wish to remain relevant, they must instill in students a love for learning. A learning organization requires employees to share knowledge and be committed to continuous learning. It is difficult to imagine a creative worker who has no interest in acquiring new knowledge and is opposed to change.

Conclusion

In 2009, 90% of executives identified *organizational agility* as a "core differentiator in today's rapidly changing business environment" (Economist Intelligent Unit, 2009). Researchers at McKinsey & Company have studied organizational health and found that firms have to be flexible and move quickly to take advantage of opportunities or to deal with threats because of the rapidly changing business environment (Bazigos, Smet & Gagnon, 2015).

The coronavirus pandemic sheds light on what happens to a firm or industry that lacks organizational agility and responds slowly to threats. Industries devastated by COVID-19 include retail, gambling, hotel, airline, movie theater, live sports, cruise, film production, automaking, theme parks, gyms, construction, and transportation (Suneson, 2020). Firms that survive will be resilient, agile, and innovative and need agile leaders who have the courage and ability to change course quickly and make fast decisions (Forbes Coaches Council, 2018; Friedman & Lewis, 2021; Friedman & Lewis, 2014; Raphan & Friedman, 2014; Schwab, 2016).

Firms today cannot expect to be around forever. One only needs to examine the list of Fortune 500 companies from 1955 and then look at it from 2019. In those 64 years, only 52 companies appear on both lists. About 90% of the Fortune 500 companies from 1955 have gone bankrupt, merged, or shrunk. The average tenure on the Fortune 500 list was 33 years in 1965, 20 years in 1990, and it will likely shrink to 14 years by 2026 (Perry, 2019). Every Fortune 500 CEO should know that there is a 50% chance that their firm might disappear from the list during the next ten years.

In the military, "fighting the last war" has become shorthand to describe the folly of using old strategies when fighting current battles. Just as the Industrial Revolution changed the way wars are fought, digitization has changed the rules of competition. Organizations that succeed will have a flexible, futuristic leader who understands the importance of agility and speed and knows when to change course. Suppose higher education is going to prepare its graduates for the workforce properly. In that case, it will have to ensure that the curriculum teaches students critical thinking, creative thinking, how to work in teams, respect for diversity, active listening, and to have a passion for learning.

References

Argyris, C. & Schoen, D. (1996). *Organizational learning II: Theory method, and practice*. Reading, MA: Addison-Wesley.

Bazigos, M., De Smet, A. & Gagnon, C. (2015, December). Why agility pays. *McKinsey & Company*. Retrieved from http://www.mckinsey.com/insights/organization/why agility pays

Brosseau, D., Ebrahim, S., Handscomb, C. & Thaker S. (2019, May 10). The journey to an agile organization. *McKinsey.com*. Retrieved from https://www.mckinsey.com/business-functions/organization/our-insights/the-journey-to-an-agile-organization

Burt, C. (2022, June 2). How well does your college encourage students to freely express views? *University Business*. Retrieved from https://universitybusiness.com/how-well-does-your-college-encourage-students-to-freely-express-views/

Business Insurance Staff (2011). 10 worst business decisions in history. *Businessinsurance.org*. Retrieved from http://www.businessinsurance.org/10-worst-business-decisions-in-history/

Care, E. & Luo, R. (2016). Assessment of transversal competencies: policy and practice in the Asia-Pacific region. *UNESCO.* Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000246590

Cotton, (2019, March 21). VUCA: The term you need to know. *Business Leader*. Retrieved from https://www.businessleader.co.uk/vuca-the-term-you-need-to-know/62095/

Council of State Governments (2015 April). America's knowledge economy: A state-by-state review. *Elsevier Report.* Retrieved from http://www.csg.org/programs/knowledgeeconomy/Elsevier_Report_2015.pdf

Davis, L. C. (2018, December 17). Creative teaching and teaching creativity: How to foster creativity in the classroom. *Psych Learning Curve*. Retrieved from http://psychlearningcurve.org/creative-teaching-and-teaching-creativity-how-to-foster-creativity-in-the-classroom/

Dyer, L. (2011). *Critical thinking for business students*. West Concord, Ontario: Captus Press

Economist Intelligent Unit (2009 March). Organisational agility: How business can survive and thrive in turbulent times. *Economist*. Retrieved from http://www.emc.com/collateral/leadership/organisational-agility-230309.pdf

Edwards, Richard (1999). The academic department: How does it fit into the university reform agenda? *Change 31*, 17-27.

Ennis, R. H. (1997). Incorporating critical thinking into the curriculum: An introduction to some basic issues. *Inquiry: Critical Thinking Across the Disciplines*, 16(3), Spring, 1-9.

Forbes Coaches Council (2018). What does it mean to be an agile leader? *Forbes*. Retrieved from https://www.forbes.com/sites/forbescoachescouncil/2018/06/29/what-does-it-mean-to-be-an-agile-leader

Friedman, H. H. (2017, June 30). Cognitive biases that interfere with critical thinking and scientific reasoning: A course module. Available at SSRN: https://ssrn.com/abstract=2958800 or http://dx.doi.org/10.2139/ssrn.2958800

Friedman, H. H. (2020). Organization agility, visionary leadership in the age of VUCA. Europe PMC. Retrieved from https://europepmc.org/article/ppr/ppr240426

Friedman, H. H. & Lewis, B. J. (2014). A new kind of C.E.O. for the global information age. *Business Quest*. Retrieved from http://www.westga.edu/~bquest/2014/ceo2014.pdf

Friedman, H. H. & Lewis, B. J. (2021). The importance of organizational resilience in the digital age. *Academia Letters*, July. Available at: https://www.academia.edu/49926596/The_Importance_of_Organizational_Resilience_in_the_Digital_Age

Friedman, H. H., Frankel, M., & Friedman, L. W. (2016, January 13). Teaching statistics with scenarios to foster critical thinking. Available at SSRN: http://ssrn.com/abstract=2715044 or http://dx.doi.org/10.2139/ssrn.2715044

Friedman, H. H., Friedman, L. W. & Leverton, C. (2016). Increase diversity to boost creativity and enhance problem-solving. *Psychosociological Issues in Human Resource Management*, 4(2), 7-33.

Friedman, H. H., Friedman, L.W., & Pollack, S. (2005). Transforming a university from a teaching organization to a learning organization, *Review of Business*, 26(3), 31-35.

Gladwell, M. (2011, May 16). Creation myth: Xerox, PARC, Apple, and the truth about innovation. *New Yorker*. Retrieved from http://www.newyorker.com/reporting/2011/05/16/110516fa_fact_gladwell

Hirt, M. & Wilmott, P. (2014). Strategic principles for competing in the digital age. *McKinsey Quarterly*, May. Retrieved from http://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/strategic-principles-for-competing-in-the-digital-age

Jurberg, A. (2021). 7 of the biggest business mistakes in modern history and what we can learn from them. *Entrepreneur's Handbook*. Retrieved from https://entrepreneurshandbook.co/7-of-the-biggest-business-mistakes-in-modern-history-and-what-we-can-learn-from-them-c08b8193bb21

Klein, Julie T. (1996). *Crossing Boundaries: Knowledge, Disciplinarities, and Interdisciplinarities*. Charlottesville, VA: University Press of Virginia.

Lee, N. (2018, January 18). The simple secret to groundbreaking science: work together. *World Economic Forum.* Retrieved from https://www.weforum.org/agenda/2018/01/the-secret-of-groundbreaking-science-global-collaboration/

Lewis, P. (2016, May 27). Is knowledge doubling or halving. *Worldnetdaily.com*. Retrieved from http://www.wnd.com/2016/05/is-knowledge-doubling-or-halving/

Lindner, M. (2008, March 10). The top 10 business blunders. *Forbes*. Retrieved from http://www.forbes.com/2008/03/10/ford-microsoft-xerox-ent-manage-cx_ml_0310blunder_slide.html

Lohr, S. & Tejada, C. (2018, February 1). After era that made it a verb, Xerox, in a sale, is past tense. *New York Times*, A1, A11.

Morgan, J. (2015, July 6). The 5 types of organizational structures: part 1, the hierarchy. *Forbes*. Retrieved from http://www.forbes.com/sites/jacobmorgan/2015/07/06/the-5-types-of-organizational-structures-part-1-the-hierarchy/

Nazar, J. (2013, November 7). The biggest business blunders in history. *Business Insider*. Retrieved from http://www.businessinsider.com/the-biggest-business-blunders-in-history-2013-11

Paul, P. (2022, September 19). When diversity isn't the right kind. *New York Times,* A19.

Perry, M. J. (2019, May 24). Comparing 1955's Fortune 500 to 2019's Fortune 500. *Foundation for Economic Education*. Retrieved from https://fee.org/articles/comparing-1955s-fortune-500-to-2019s-fortune-500/

Raphan, M. & Friedman, H. H. (2014). Adaptivity, Creativity, and Diversification in the Age of Convergence. *Business Quest*. Available at: http://www.westga.edu/~bquest/2014/convergence2014.pdf

Ray, A. (2020, October 22). Human knowledge is doubling every 12 hours. *LinkedIn*. Retrieved from https://www.linkedin.com/pulse/human-knowledge-doubling-every-12-hours-amitabh-ray/

Robson, D. (2020, June 9). How to avoid the 'competency trap.' *B.B.C. Worklife.* Retrieved from https://www.bbc.com/worklife/article/20200608-what-is-the-competency-trap

Roesler, P. (2015, April 20). 8 of the biggest business mistakes in history. *Inc.com*. Retrieved from http://www.inc.com/peter-roesler/8-of-the-biggest-business-mistakes-in-history.html

Schlueter, J. (2016, June 7). Higher ed's biggest gamble. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/views/2016/06/07/can-colleges-truly-teach-critical-thinking-skills-essay

Schwab, K. (2016, January 14). The fourth industrial revolution: What it means, how to respond. *WEForum.org*. Retrieved from https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/

Schwantes, M. (2022). How can you be sure someone has exceptional leadership skills? It comes down to 2 words. *Inc.com*. Retrieved from https://www.inc.com/marcel-schwantes/how-can-you-be-sure-someone-has-exceptional-leadership-skills-it-comes-down-to-2-words.html

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. New York: Doubleday.

Shontell, A. (2011, January 19). The amazing story of how Steve Jobs took Apple from near bankruptcy to billions in 13 years. *Business Insider.* Retrieved from https://www.businessinsider.com/how-steve-jobs-took-apple-from-near-bankruptcy-to-billions-in-13-years-2011-1

Stone, B. (2021, July 4). The Bezos paradox. *New York Times*, SR3.

Suneson, G. (2020, March 20). Industries hit hardest by coronavirus in the U.S. include retail, transportation, and travel. *24/7 Wall Street*. Retrieved from https://www.usatoday.com/story/money/2020/03/20/us-industries-being-devastated-by-the-coronavirus-travel-hotels-food/111431804/

Tett, G. (2015). *The silo effect: The peril of expertise and the promise of breaking down barriers.*

Weinberger, M. & Hartmans, A. (2020, February 24). Steve Jobs would have been 65 on Monday. Here's how the late Apple C.E.O. saved the company from disaster and set it on the path to a $1 trillion valuation. *Business Insider*. Retrieved from https://www.businessinsider.com/steve-jobs-apple-photos-2017-1

OSS Project Community (OSSPC) Ambidexterity: Its Social Network Antecedents and Community Performance Impact

Abstract

This study is to conceptualize a model about the social network antecedents of both the balanced and combined ambidexterity of OSS project community (OSSPC) and the impact of interactions between balanced and combined ambidexterity on OSSPC performance. It intends to advance the OSS research by combining both the social network and the sustained participation perspectives. These two perspectives together represent information and knowledge exchanges and trust building and manifest OSSPC ambidexterity. Further, empirical testing procedures using time series data are also proposed to lay the foundation for advancing the study. The proposed study may explain the potential advantages of the OSS development model over the proprietary model of software development.

Introduction

Open-source software development has been a valid and powerful model of software development for several decades. There are many very successful OSS projects including such as Linux, Apache Server, Apache Spark, Apache Hadoop, etc. Many studies have been conducted to investigate the reasons for the success of the OSS development model (e.g., Shi and Sun, 2021; Fang and Neufeld, 2009; Singh, et al., 2011; Temizkan and Kumar, 2015). Among these studies, there are two research streams including the social network perspective and the sustained participation perspective. The current study is to advance the OSS development research by integrating both of these two research streams in the sense of treating social network properties as the antecedents and placing the OSSPC ambidexterity as the mediating mechanism to enhance OSSPC performance.

Singh, et al. (2011) introduced the social network perspective to the research of open source project success. Their definition of success for open-source projects is knowledge creation in terms of the number of CVS commits, representing completed modification requests. Their social network perspective focuses on the structural aspect of the OSSPC network, measured as the project level internal cohesion and external cohesion. They included control variables such as technology diversity (i.e., the difference between the focal project and other projects which share at least one developer), direct ties and indirect ties to developers not in the focal project, project human capital and ability (i.e., project size and pre-sample project commits), user input, market potential, project age, and a few other characteristics. While they generated interesting findings, their research did not distinguish between exploitative and explorative activities in software development. To advance this stream of research, Temizkan and Kumar (2015), following Singh, et al. (2011), investigated the issue in more refined details. They separate exploitation social network (i.e., the social network of those developers involved in patch development) from the exploration network (i.e., the social network of those developers involved in new feature requests) and their project success is defined as the knowledge creation in terms of the number of files created for patch development and feature requests respectively. They measured social networks not only using characteristics such as internal cohesion and external cohesion, but also the network location and the macro level measure of network decomposition. They found that social networks of exploitation and exploration are

different, and these social networks have varied impacts on exploitation and exploration. More interestingly, their study incorporated the concept of ambidextrous developers, representing those who work on both exploitative and explorative tasks.

Based on the legitimate peripheral participation theory, Fang and Neufeld (2009), through a case study of the phpmyadmin community, developed a sustained participation model indicating the impacts of situated learning and identity construction on sustained OSS community participation. Shi and Sun (2021), based on the integrative model of trust, empirically tested the continuous overall positive interacting effects between community trust and community citizenship behaviors (as participating activities). These sustained participation-oriented studies clearly demonstrate the importance of community participation in explaining OSS project performance and success. The current study is to integrate these two streams of research by treating OSSPC social network properties as antecedents of OSSPC ambidexterity (manifesting the nature of various participation activities) for enhanced OSSPC performance.

Theory Development

Conceptual Model

First, this study will investigate how the social network of an OSS project community affects the OSSPC ambidexterity and how the ambidexterity influences OSSPC performance *over time with time series data*. As such, the current study supplements previous studies from an evolving perspective. Second, this study will incorporate more fine-tuned ambidexterity measurement including the balanced ambidexterity (i.e., |exploitation-exploration|) and the combined ambidexterity (i.e., exploitation * exploration) (Cao, et al., 2009). Cao et al. (2009) found that the interactions between these two ambidexterity constructs have positive impacts on organizational performance and no studies have investigated these constructs in the context of an OSS project community. *So, the key research question of this study is how the findings in the proprietary firms regarding ambidexterity and performance will look like in the context of OSSPCs* as OSSPCs can only be roughly treated as virtual organizations without legal employment relationships, clear organization structure, and strictly enforceable rewarding systems. More specifically, the following questions will be investigated:

1) *How does the social network of an OSSPC evolve over time?*
2) *How do the balanced ambidexterity and combined ambidexterity of an OSSPC evolve over time?*
3) *How does the evolving social network of an OSSPC affect its (balanced and combined) ambidexterity over time?*
4) *How does the balanced ambidexterity and the combined ambidexterity of an OSSPC jointly affect its performance over time?*

Answers to these questions enhance our understanding of the interactions between exploitative and explorative activities in the context of an OSSPC and can guide the OSS project committee and firm sponsors manage the OSSPC for its success and continuous prosperity. The following Figure 1 summarizes these research questions in a conceptual model.

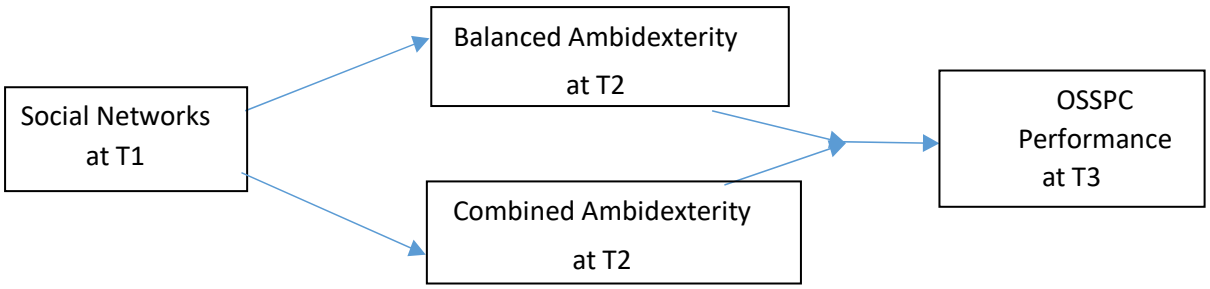


Figure 1: Conceptual Model

Impact of OSSPC Ambidexterity on OSSPC Performance

OSSPC ambidexterity is about the balance between exploitative and explorative activities in the community to develop software systems. Exploitation is about refinement and incremental improvement and exploration is about new future development and application of new technologies. Cao, et al. (2009) developed the constructs of balanced ambidexterity (i.e., BD=|exploitation-exploration|) and combined ambidexterity (i.e., CD=exploitation*exploration) in the context of organizations and found that instead of exploitation, exploration, balanced ambidexterity, or combined ambidexterity themselves, only the interaction between balanced ambidexterity and combined ambidexterity (BD*CD) has a positive impact on firm performance. This finding demonstrates that for firms to be successful they need to balance the allocation of resources for exploitative and explorative activities. This is not only in the sense that the numbers of exploitative and explorative activities should be relatively equal, but also in the sense that the more of both exploitative and explorative activities, the better for firm performance.

In the context of an OSSPC, since it can be roughly treated as a virtual organization, the interactions between balanced ambidexterity and combined ambidexterity could have a positive impact on OSSPC performance. More specifically, OSSPC exploitative activities eliminate software bugs and refine the software performance to make the existing version of a software system more stable and usable in a production environment. OSSPC exploration activities, on the other hand, advance the software system in the sense that new features are developed over time and radical performance improvement may be made with better modularization and new technologies. These new features may attract more potential users, better modularization makes the development more suitable for the OSS model as modularization facilitates the global participation of developers, and new technologies may help the software system meet special performance demands. Specifically, with a high level of balanced ambidexterity, the OSSPC software is developed and maintained for meeting both the short-term user needs and its long-term viability in the competitive software market. With a high level of combined ambidexterity, more exploitative and explorative activities are conducted to better satisfy both the production use of the software and the experimental and future-oriented evaluation of the software. Consequently, it is hypothesized that:

H1: The interactions between balanced ambidexterity and combined ambidexterity in an OSSPC has a positive impact on the OSSPC performance

Social Network of an OSSPC and its Impact on the OSSPC Ambidexterity

OSSPC members form social relationships through forum discussions, code developments, and other activities. Measures such as internal cohesion, network location, network decomposition are used to describe social networks in existing OSSPC studies. Internal cohesion represents the strength of ties among project members. Internal cohesion facilitates social exchanges and nurtures trust development within an OSSPC, motivating information and knowledge sharing and collaborative efforts. However, internal cohesion at a high level may prohibit participants from creative thinking for product innovations. Singh et al. (2011) indeed found that only a moderate level of internal cohesion is better for OSS project knowledge creation. In a more refined detail, Temizkan and Kumar (2015) found that internal cohesion relates more positively to the patch development (i.e., exploitation) than the feature request (i.e., explorations). Consequently, it is hypothesized that:

H2a: A moderate level of internal cohesion is better for OSSPC balanced ambidexterity over time.
H2b: A moderate level of internal cohesion is better for OSSPC combined ambidexterity over time.

Network location represents the centrality of OSSPC members. An OSSPC member with a high level of centrality is positioned at a relatively central location in a social network, controlling information and knowledge flow and having a more complete picture of the structure of the social network, facilitating task completion. Temizkan and Kumar (2015) found that centrality relates more to the patch development (i.e., exploitation) than the feature request (i.e., explorations). However, too much centrality may overload OSSPC members with redundant information, spread their energy too thin, and produce unbalanced exploitation and exploration. Consequently, it is hypothesized that:

H3a: A moderate level of centrality is better for OSSPC balanced ambidexterity over time.
H3b: A moderate level of centrality is better for OSSPC combined ambidexterity over time.

Network decomposition represents the connectedness of subgroups within the social network. Within each sub-group, all OSSPC members are reachable with each other, facilitating information and knowledge exchanges. OSSPC members from different subgroups are not well connected. Like internal cohesion and centrality, with a high level of network decomposition, while exploitation activities are encouraged and facilitated, necessary exploration activities that are deviant from sub-group norms and standards may be reduced and it may prohibit novel ideas from being generated and applied for software innovations. Consequently, it is hypothesized that:

H4a: A moderate level of network decomposition is better for OSSPC balanced ambidexterity over time.
H4b: A moderate level of network decomposition is better for OSSPC combined ambidexterity over time.

Research Method

Data Source

Both Singh, et al. (2011) and Temizkan and Kumar (2015) used data from the sourceforge.com across many projects. The proposed research, instead, will use the social network data embedded

in OSSPC forums. Thread discussions in community forums may represent OSS development tasks (such as bug clarification, bug fixing, new feature request, solution proposal, user training and assistance, etc.) and the process of the construction of the social network in an OSSPC. For the ambidexterity data, OSSPC commits (for both explorative and exploitative activities) will be obtained for calculating the ambidexterity indices. OSSPC performance can be measured as product downloads, forum views, and community growth.

Past studies (using all projects from sourceforge.com) count ties among developers by checking whether they worked on the same project previously. The potential problem of this counting method in measuring social networks is that two developers who were involved in the same project previously may never have a direct communication with each other. While they surely have some common background (such as using the same programming language and APIs and the general embedded business processes), they may never socialize with each other and truly establish a personal tie. This lack of personal tie in social networks used in previous studies may prevent important findings from being identified.

Data Preparation

This paper focuses on the impact of social networks on the OSSPC ambidexterity and the impact of ambidexterity on OSSPC performance over time. For every 3 or 6 months, the social network (embedded in thread data) of an OSSPC is constructed and data regarding the OSSPC exploitative/explorative activities and its performance are collected.

Construction of Thread-Based Social Networks

The social network of an OSSPC can be constructed based on forum thread discussions. Thread discussions are rich in content and represent the socialization processes. OSSPC members involved in a thread discussion form an ad-hoc team and have opportunities to establish ties through exchanging information and knowledge and collaborating on self-selected project tasks. A complete social network for an OSS community can be constructed by connecting members based on whether they participate in the same thread discussions. Obviously, as time goes on, there are more thread discussions, indicating more opportunities for members to form social ties and reinforce existing ties.

Measurement

OSSPC performance: OSSPC performance consists of community growth, product downloads and forum page views. As the size of the community is growing over time, the OSSPC has a higher potential to produce better software products as participation is the key to the success of an OSSPC. Product downloads measure the number of potential users who are testing the software and the number of developers who are interested in learning the system and getting involved in developing the system. Forum page views represent the general interests in the community and the popularity of the software in the world. These measures, while overlapped to some degree, together depict the

whole picture of OSSPC performance.

Interactions between balanced ambidexterity and combined ambidexterity: The numbers of exploitative activities and explorative activities are counted for every 3 or 6 months based on community commit data. Then both the balanced ambidexterity (as the absolute value of the difference between the number of exploitative activities and the number of explorative activities) and the combined ambidexterity (as the product of the number of exploitative activities and the number of explorative activities) are calculated. Finally, the interaction between balanced and combined ambidexterity is measured as the product of these two ambidexterity indices.

Internal Cohesion: Repeated ties over different thread discussions are used to capture an OSSPC's internal cohesion. The total number of threads in which each pair of OSSPC members has participated are counted. Then this number is divided by the total number of pairs that exist in the OSSPC to represent the repeat ties in the community. Obviously, a high score represents a higher average number of repeated ties, indicating a high possibility of tight internal cohesion.

Network Centrality: For each OSSPC member, the degree of centrality is calculated based on Freeman (1978). Then the average of all OSSPC members' degree of centralities will be used as the measure of the centrality for the OSSPC itself. A high score of this measure indicates the OSSPC is more centralized.

Network Decomposition: For the social network of an OSSPC, component analysis (Nooy, et al., 2005; Wasserman and Frost, 1994) will be used to identify subgroups/components. Then the total number of connections within each component is counted and it is divided by the number of OSSPC members in that component (to represent the average number of connections per member in that component). Next, based on the average number of connections per member in each component, the average number of connections per component for each member is calculated to represent the network decomposition of the social network. A high value of this measure indicates that the social network has tight connectedness within each subgroup.

Analysis Procedure

First, the following notations are used for brevity:

- Internal Cohesion, Network Centrality, and Network Decomposition at time t_i are represented as IC(i), NC(i), ND(i),
- Balanced ambidexterity and combined ambidexterity at t_{i+1} are represented as BD(i+1) and CD(i+1). OSSPC performance at t_{i+2} is represented as OP(i+2).

Second, the following equations will be generated with the VAR model or other relevant time series models for hypothesis testing.

- Independent variables: IC(i)^2, NC(i)^2, ND(i)^), Dependent variable: BD(i+1),
- Independent variables: IC(i)^2, NC(i)^2, ND(i)^2), Dependent variable: CD(i+1),

- Independent variables: BD(i+1) * CD(i+1)), Dependent variable: OP(i+2).

Conclusion

The current study is intended to supplement extant research by integrating the social network perspective and the sustained participation perspective. Time series analysis is unique in helping understand the evolution of social networks and their impacts on OSSPC ambidexterity. The incorporation of both the balanced ambidexterity and combined ambidexterity likely produces refined and detailed insights into OSS ambidexterity and its impact of OSSPC performance. This proposed study may help explain potential advantages of the OSS development model over the proprietary model of software development.

References

1. Singh, P. V., Tan, Y., & Mookerjee, V. (2011). Network effects: The influence of structural social capital on open source project success.
2. Cao, Q., Gedajlovic, E., & Zhang, H. (2009). Unpacking organizational ambidexterity: Dimensions, contingencies, and synergistic effects. *Organization Science*, *20*(4), 781-796.
3. Temizkan, O., & Kumar, R. L. (2015). Exploitation and exploration networks in open source software development: an artifact-level analysis. *Journal of Management Information Systems*, *32*(1), 116-150
4. Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, *1*(3), 215-239.
5. Nooy, W.D.; Mrvar, A.; and Batagelj, V. Exploratory Network Analysis with Pajek. Cambridge: Cambridge University Press, 2005.
6. Wasserman, S., and Frost, K. Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press, 1994.
7. Shi, Z. and Sun, H., 2021. Sustained Participation in Open-Source Software Project Communities. Journal of Computer Information Systems, pp.1-14.
8. Fang, Y. and Neufeld, D., 2009. Understanding sustained participation in open source software projects. Journal of Management Information Systems, 25(4), pp.9-50.

Legal, Ethical, and Social Issues

CORPORATE HYPOCRISY AND NEGATIVE EXTERNALITIES IN BANKRUPTCY CARVEOUTS: THE TEXAS TWO-STEP

INTRODUCTION

Business ethics as a discipline has proliferated business school curriculum and corporate governance documents since the turn of the century. The heavy absorption of ethical training, while obviously pre-dating 2000, came about as a reaction to the corporate scandals during this time frame, most notably Enron and Arthur Anderson. It is debatable if corporate America has become more ethical or just better at virtue signaling its collective ethical efforts. In the first two decades of the twenty first century, almost all large, publicly traded firms have codified ethical statements in Codes of Conduct (COC) or Corporate Social Responsibility (CSR) documents. While there have been efforts to publish and disseminate these statements, there has also been evidence that many firms have morphed their language with each other's to a point where the homogeneity of claims leads to a credibility gap in the firms' sincerity. In other words, firms have societal pressure to publish codes and use the same format and language as their peers, leading to a diminution of credibility from the lack of inter-firm differences.

Measuring ethical behavior is difficult due to the inherent subjectivity in what separates ethical behavior from either purely illegal actions or "gray area" matters. The current paper proposes to narrow this objectivity gap in studying potential unethical behavior in two ways. The first is labeled "Word-Action Reconciliation", which is how effective firms are at matching their words, be they published or spoken, with their subsequent actions. The second is attempting to measure the externalities that result from a firm's actions. In order to study this two-pronged approach to ethical behavior, I use a multiple case study approach with firms that have used a controversial legal maneuver called the Texas Two-Step. The Texas Two-Step entails a large firm

executing a spin-off in the face of numerous product liability claims. The first step is in creating the spin-off, which becomes a separate legal entity from the original firm. The second step involves packing the new spin off with all legal claims from the product liability event for the sole purpose of having the newly spun off entity file for bankruptcy protection while leaving the original firm out of the bankruptcy system.

The Texas Two Step has been deemed legal by at least one appellate court in the United States and, therefore, some may argue that the firm has the right to form such a spun off subsidiary. However, ethically, firms may have obligations that supersede legal thresholds if the actions taken violate societal principles or, more importantly, their own stated principles. Since the goal of the Texas Two-Step is to shield the original firm from liability arising from legal claims by diverting these claims in an artificial manner (or at least quasi-artificial) to an entity that exists for the sole purpose of filing bankruptcy, this is an appropriate phenomenon to study the two-pronged approach to ethical behavior as described above.

The rest of the paper is as follows. The next section is the Theory Section, which will explore the underpinnings for ethical objectivity. Following this is the Methods Section, which describes the case study approach, pertinent information on the process for reorganization under Chapter 11 in the U.S. Bankruptcy system, and a description of the Texas Two Step legal strategy. After this is the empirical section that contains three case studies on large, U.S. firms that have attempted the novel, and controversial, Texas Two Step namely Georgia-Pacific, Johnson & Johnson and 3M. Finally, I address the analysis of the findings and how these findings mesh with theory in the Discussion Section.

THEORY

Action-Word Reconciliation

Objectivity in ethics is difficult to reach consensus on since the nature of ethics is rooted in debatable premises. While objectivity is difficult, one area of objectivity in business ethics can be found in the hypocrisy literature. Hypocrisy is "feigning to be one which one is not" (Merriam-Webster Dictionary 2022) or "the practice of claiming to have moral standards or beliefs to which one's own behavior does not conform" (dictionary.com). While the definitions are individual and human in nature, corporations can also be hypocritical when the entity itself claims or publicizes beliefs that it holds true. That is, while individuals within the firm have their own set of standards by which they live, the firm itself can also have moral standards by which it claims to operate (i.e. live) by.

Corporate hypocrisy is defined here as the mismatch between a firm's words (written or verbalized) and its subsequent action whereby it does not behave in a manner that conforms to its words. Following Wagner, Lutz and Weitz (2009), however, hypocrisy may also be defined in the opposing order, that is when a firm acts in one way and then defines its ethics differently. Corporations claim their moral standards or virtues in numerous settings including, but not limited to, Codes of Conduct (COC), Corporate Social Responsibility (CSR) documents, press releases, and official interviews given by corporate officers when the officers are giving statements in their professional (as opposed to personal) capacities. An important part of the statements made in determining if hypocrisy is present or not is in the voluntary nature of the written or spoken words. Some documents published by corporations are involuntary, such as securities filings mandated by regulators such as the Securities and Exchange Commission (SEC). For a true hypocritical analysis, however, one must look at those documents that are non-compulsory. COCs and CSRs

are voluntary documents that firms do not have to compile by law and are usually proactive in their language as to which virtues the entity conducts itself.

The literature on corporate hypocrisy is scant but reveals some common notions that serve as a platform for further scholarly development. Jauernig, Uhl and Valentinov (2021) explain that corporations are hypocritical if they "…pretend to be socially responsible" while Wagner, Lutz and Weitz (2009) think that when a firm "…appears something that it is not", then hypocrisy ensues. In addition to defining the ethical barriers of corporate hypocrisy, some studies have utilized qualitative and quantitative approaches.

One sub-stream of the corporate hypocrisy literature focuses on stakeholders' perceptions of a firm's hypocrisy. Scheidler, Edinger-Schons, Spanjol and Wieseke (2019) found that employees perceived firms to be hypocritical when their CSR initiatives were focused more on external stakeholders than internal. Jauernig et. al. (2021) used an experimental methodology in a sample of 617 adults who evaluated several vignettes concerning individual and corporate behavior. Subjects found corporate hypocrisy more prevalent than individual hypocrisy in scenarios that were illegal in nature and where the firm was part of an NGO that claimed to be against similar behaviors. Losada-Otalora and Alkire (2021) found that bank customers who benefitted from CSR activities were less prone to perceive their banking institutions as hypocritical. Conversely, Arli, Grace, Palmer and Pham (2017) studied the relationship in the opposite direction and found that customers' corporate hypocrisy perceptions negatively influenced the believability of the focal firm's CSR statements as well as its reputation.

Externality Theory

Externality Theory explains the effects that multi-actor events have on those outside of the direct event. While externalities may be positive or negative, the bulk of academic research on the subject concerns the latter as negative effects may be more consequential to study. Negative externalities are, therefore, negative effects (i.e. costs) imposed upon a societal actor from an interaction between two or more others. In an interaction between Actor A and B, a negative externality occurs when deleterious harm is borne by Actor C but where Actor C was not part of the direct interaction. Negative externalities can include more than one outside actor (Actors D, E, etc) and can have effects that have multi-layer reverberations (secondary and tertiary effects). Published work on negative externalities is quite robust as the notion has been present in the literature for up almost two centuries.[1]

Much of the externality literature focuses on negative effects of spillovers. A number of papers have attempted to quantify location-based asset price diminution for real estate (Epstein 2018) that is located near places such as U.S. airports (Scotti, Dresner, Martini and Yu 2014), railyards and similar freight (Demir, Huang, Scholts and Van Woensel 2015; Diao, Qin and Sing 2016), and wind turbines (Kreke and Zerrahn 2017). In addition to real estate pricing, some externality literature models the negative effects of CSR efforts on marketing effectiveness (Grolleau, Ibanex and Lavioe 2016) showing a tradeoff between CSR and profit-based marketing initiatives. Other work has focused on contagion effects, such as the propensity of municipal bankruptcies when geographically-correlated bankruptcies take place (Yang 2019). Finally, and based corporate cronyism, some scholars have modeled deleterious effects to parts of society as

[1] The original externality theorist has been debated by scholars with some claiming that Henry Sidgwick formulated the notion in the mid-1800s and others siding with Arthur Pigou in the early 1900s. The origin for the purposes of this paper is inconsequential since either would constitute the idea as being long-running and engrained in the literature (O'Donnell 1976).

firms or industries capture government officials (Enderwick 2005; Manish and Sutter 2016; Brown 2018, 2022).

Kantian Ethics

Business ethicists have debated the motives behind corporate decisions that address ethical considerations. Purists follow a Kantian argument where intentions are paramount to outcomes. This contrasts with a more Hobbesian approach where the outcome is the ethical measurement and the underlying motive is a secondary consideration. Under the first scheme, a firm must have good intentions with respect to its actions regardless if the outcome to such actions results in a betterment to its efforts. Under the second scheme, as long as the outcome benefits the target constituency, then the methods to get there are irrelevant. This boils down to the "means justify the ends" versus the "ends justify the means" (Wootton 2018).

However, since firms implicitly take a stance on this debate, then it can be integrated into the current theoretical framework. The signaling and messaging derived from corporate documents leads to the conclusion that firms are making COD and CSR efforts for the good of society. In other words, firms generally imply that they are Kantian and not Hobbesian since subscribing to the latter would appear disingenuous. While most stop short of labeling these actions altruistic, altruism is implied through the lack of shareholder-centric reasoning provided in corporate messaging or codification.

METHODS

Case Methodology

Case studies can be an effective methodology, especially when a phenomenon is new and, therefore, without academic precedent and/or the sample of the phenomenon is relatively low

(Gibbert, Ruigrok and Wicki 2008). According to Gibbert et.al. (2008), four attributes contribute to rigor in the case study methodology, namely (i) internal validity, (ii) construct validity, (iii) external validity and (iv) reliability. Internal validity addresses causal relationships between variables where the relationship between variables and results are logically linked. Construct validity is defined as a case study that studies what it claims to study at the outset. External validity entails the generalizability of the study's findings. External validity is concerned with the analytical, as opposed to a statistical, generalizability. Finally, reliability means that the study has the ability to be replicated, which focuses on the access to the qualitative, empirical data a study reports on. The current research was completed with these four important criteria in mind and arguments of compliance to these standards are made in the Discussion Section.

Since the Texas Two Step is a relatively new legal strategy, cases were chosen non-randomly by researching the few firms that have utilized it. These include Georgia-Pacific, Johnson & Johnson and 3M, which are all multi-national publicly traded firms domiciled in the United States. In each case, I compiled documents concerning the litigation that preempted the bankruptcy strategy, the initiation of the Texas Two Step, the bankruptcy petitions filed in the U.S. federal court system and documents or filings related to judicial determinations in the U.S. federal court system. Documents were extracted from the business press (Wall Street Journal, Financial Times, CNBC, etc), federal court records at the Public Access to Court Electronic Records (PACER), and court filings from bankruptcy administrators such as Prime Clerk. All told, more than 1,000 pages of documents were compiled to analyze the qualitative empirics for the case studies that follow.

Chapter 11 Reorganization in the U.S. Bankruptcy System

Before delving into the cases *per se*, however, an explanation of the process of reorganization in the U.S. Bankruptcy Courts under Chapter 11 is warranted. This will be followed by an explanation of the corporate restructuring needed for the Texas Two Step.

Chapter 11 Reorganization Primer

U.S. Code, Title 11 is the part of the federal code that addresses bankruptcy, both individual, corporate and governmental. There are a number of "chapters" that can be filed depending on what type of debtor is involved and these chapters are named for the chapter of Title 11 that they fall under. Chapter 7 is synonymous with liquidation and can be utilized by organizations and individuals. On the other hand, Chapter 11 is synonymous with reorganization. Chapter 11 reorganizations are typically completed by corporations, but individuals may also utilize this chapter if they do not qualify for relief under Chapter 13 (individuals only). The three other chapters (9, 12 and 15) are idiosyncratic as they are for bankruptcies filed by municipalities (Chapter 9), farmers and fishermen (Chapter 12) and international firms (Chapter 15). This paper only focuses on Chapter 11, which is where corporations try to become more efficient and remain a going concern (i.e. reorganize) post-bankruptcy.[2]

Chapter 11 begins with the filing of a bankruptcy petition, which places an automatic stay on creditor collection of debtor assets. The automatic stay provision allows an orderly planning process around the debtor firm's asset and liability reorganization. For 180 days following the filing date, the debtor firm has the exclusive right to create and disseminate a bankruptcy plan ("The Plan") that addresses the different categories of creditors that have unmet obligations as of the date of the filing. During this time, creditor committees are being formed, which group

[2] See the U.S. Federal Code at www.uscode.house.gov

creditors into like claims. Common creditor committees are secured creditors, which are those creditors owed an obligation on specific real property and will each become its own committee, unsecured creditors, which are those creditors owed an obligation but which are not secured against any real property and government, which may be owed taxes. The plan must be voted on by all of the creditor committees, but they are not all equally important to the bankruptcy court.

To understand the salience of committees, one must understand the difference not between secured and unsecured, but that between (i) unsecured and impaired versus (ii) unsecured and unimpaired. Impairment is possibly the most important factor in gaining approval of the plan, as it is the latter committee that must vote for the plan (or at least one of these committees if there are more than one). An unsecured creditor is unimpaired if its contractual terms are not changed due to the bankruptcy. An example is a consumer who holds a warranty on a product when the product-firm files for bankruptcy. If the firms states that it will honor the claims post-bankruptcy, then this unsecured creditor has not been impaired by the filing. More important are those unsecured creditors who are impaired. An unsecured creditor is impaired if any part of its contract with the debtor firm will change as a result of the bankruptcy filing (11 U.S. Code § 1124). While most would intuit this to be a reduction in the amount owed to the creditor, it could also be any other aspect (i.e. interest rate, payment terms, warranty cancellation, etc.) that places the creditor in a worse post-filing position as a result of the plan. In order for a plan to be approved, which allows the debtor firm to exit bankruptcy as a more efficient going concern, a majority in number and two-thirds in amount of that class must vote for the plan. These thresholds are of those members that vote and not of the total in that class (Friedland 2005).

Creditor and Equity Outcomes in Corporate Chapter 11 Filings

At the heart of the question in the Texas Two Step is the pre-bankruptcy litigation against firms for mass tort product liability claims. In the previous discussion, unsecured creditors whose claims are impaired were the focus in reorganizing a corporation. An important unsecured creditor class is comprised of plaintiffs whose litigation is initiated prior to the bankruptcy filing but whose trials and verdicts were stayed by the bankruptcy filing due to the automatic stay provision. In many instances, the proximate cause of the bankruptcy filing itself was due to the numerous claims that a common product deficiency causes against the same entity. Since bankruptcy stays all ongoing litigation, it acts as a mechanism to "take a breath" and see if there is a better, collective outcome for all plaintiffs and for the debtor-defendant firm as well. In other words, if the firm faces so many lawsuits that it must defend against, then it may liquidate leading the firm to cease to exist and the plaintiffs' claims to cease to exist along with it. In a Chapter 11 reorganization, there may exist a better joint outcome between the two types of parties, the plaintiffs both individually and collectively as a group, and the firm that seeks bankruptcy protection.

At the same time, equity holders in the debtor's capital structure are in an ominous position. While shareholders are technically creditors, as residual claimants, they are last in the firm's pecking order in payment given a liquidation. In over 99 percent of corporate liquidations, equity shareholders receive nothing and in the other one percent, they receive very little with few exceptions (Freyman 2021). However, in a reorganization, equity holders typically receive no value for their pre-filing shares due to the Absolute Priority Rule (U.S. Code 1129b2), which is the legal notion that a junior class of creditors cannot retrieve any value when a class senior to them is not paid in full in the reorganization. Since shareholders are behind bond holders,

corporate lenders, and unsecured creditors, Chapter 11 is almost just as detrimental to shareholders in a reorganization as it is in a liquidation.

Mass Tort Product Liability Litigation and Shareholder Risk

The preceding sections, therefore, can be logically concluded by an explanation of the link between mass tort litigation due to product liability and the firm's shareholders in a Chapter 11 reorganization. If the litigation is serious enough, in both number of claimants and in magnitude of potential claims' verdicts, then one option for the firm is to mitigate its potential losses through a bankruptcy filing, which will facilitate the litigation and any potential payouts to plaintiffs. If done through the courts, however, this will decimate the pre-bankruptcy equity holders. Firms that are large and publicly traded are in a legal conundrum as a result since they can most likely absorb both the legal costs to the mass tort litigation and/or the payouts to any verdicts not decided in their favor. At the same time, it is tempting to lessen these potential losses by forcing the plaintiffs' claims to be negotiated in an orderly bankruptcy process. This is where the Texas Two Step comes into play as it has the potential to garner the positive outcomes from the bankruptcy process while minimizing the negative outcomes to shareholders.

The Texas Two-Step

The Texas Two-Step is a mechanism that artificially creates a spun-off entity that is comprised of all of the mass tort product liability claims that a firm faces that subsequently files for Chapter 11 bankruptcy protection. While the spin offs are quite common for large firms when they wish to exit specific product or geographic markets, in the Texas Two-Step the spinoff is a liability carrying vehicle that absorbs litigation, and the resultant potential verdicts and judgments, from

the larger firm that was the original defendant in the litigation. This legal spin off is the first step in the process.

The second step is to place the spinoff into bankruptcy to gain the protections of defending the litigation in the orderly, and favorable, court process explained above. Since the larger firm never files for bankruptcy per se, its shareholders are not subject to the deleterious effects noted, such as a 100 percent loss of their invested capital. Also part of this step, the larger firm will typically fund the litigation settlement for the bankruptcy spinoff since the new entity has no inherent assets to fund the legal settlements that may be negotiated.

This mechanism has potential legal and ethical implications. Legally, the theory behind claims priority is that the higher the potential return on risk, the greater risk of larger losses. Shareholders invest capital in firms to garner greater returns on average then they would as bondholders, who receive a generally lower payout but one which is fixed and has more priority in bankruptcy. Therefore, the Texas Two-Step's outcome is one where the shareholders of a firm that may be the cause of great harm as put forth by mass tort litigation does not share the risk of total capital loss if the litigation is insurmountable. Ethically, and here let us assume the claims are generally true, the firm escapes greater payouts by funding the spin off with a limited amount of money with the greatest payout being the collapse and liquidation of the firm to pay off the harmed plaintiffs. The shareholder(s) of the bankruptcy entity will lose all equity, but this is an artificial entity formed for the large firm. Even if the large firm is the shareholder of the new entity, its losses are smaller (fractionally smaller some would argue) than other scenarios. If allowed, the large firm pays a fixed amount to fund the spinoffs bankruptcy (i.e. legal settlements) which caps the amount payable to plaintiffs who may have otherwise been awarded exponential

damages against the large firm. This shifting of outcomes through a quasi-artificial legal maneuver poses a test of the two-prong objective ethical approach that is the theoretical focus of this paper.

CASE STUDIES

Case Study 1: Georgia-Pacific

Georgia-Pacific's history dates to 1927 when the firm then known as the Georgia Hardware Lumber Company was founded in Augusta, Georgia. After supplying the U.S. Armed Forces in World War II, the company changed its name to Georgia-Pacific Plywood & Lumber Company in 1948 and became a publicly traded firm the next year, listing on the New York Stock Exchange (NYSE). The company grew rapidly, reaching $1 Billion in annual revenue in 1968 and double this figure only five years later. In addition to general revenue growth, Georgia-Pacific became active in mergers and acquisitions as it absorbed related-industry brands such as Brawny and Quilted Northern to add to a growing list of already-owned products such as Coronet. In 2005, however, the firm became a target and was acquired by Koch Industries, which took the firm from publicly-traded to privately held. Even though it was under the Koch umbrella, Georgia-Pacific continued growing through acquisitions, eventually buying both individual mills and divisions (Oriental Strand Board facilities, Alabama Pine Mill and Alabama River Mill in 2010) as well as entire going concerns, such as Temple Inland Building Products (2013).

One early acquisition that Georgia-Pacific completed in 1965 was that of Bestwall Gypsum Company, which made joint compound from asbestos and asbestos derivatives in the years following the merger (*In Re Bestwall LLC* 2017) through the late 1970s. Georgia-Pacific made this compound even though the health threats from asbestos products were noted much earlier. For example, in a 1968 *Wall Street Journal* article, medical outcomes of men working with the

material included such diagnoses as lung cancer and general breathing issues (Kessler 1968) sometimes leading to impairment or death. While 1968 seems like early warning, it was actually a 1924 article in the *British Medical Journal* that is thought to be the first published link warning of the relationship between asbestos and serious illness (Cook 1924).

The first mesothelioma verdict related to asbestos construction was in 1971 (*Borel v. Fibreboard* 1971). In the *Borel v. Fibreboard* case, the manufacturer was held liable under the doctrine of strict liability in the United States Eastern District of Texas. The verdict was appealed and upheld by the 11th Circuit Court in 1974. Since this first trial win in 1971, it is estimated that a total of $70 Billion has been paid out in asbestos-related claims to over 700,000 individuals by over 8,000 firms (Texas State Historical Association 2020). Mesothelioma litigation began in earnest in the mid-1970s following the *Borel* verdict. For example, two class action lawsuits covering thousands of workers were filed in 1974 and 1975 against Johns-Manville Corporation in both New York and New Jersey (Author Unknown, Wall Street Journal 1975). Johns-Manville would become part of the first cohort of asbestos-related firms to file for U.S. bankruptcy protection in 1982 (Plevin *et.al.* 2016).

Georgia-Pacific would not file for bankruptcy protection like hundreds of other firms would in the 1980s and 1990s. This did not stop an onslaught of civil litigation against the firm from former employees of its subsidiaries, including Bestwall Gypsum. As cases mounted, and with some large verdicts against it, Georgia-Pacific became one of the first firms to partake in the Texas Two Step maneuver. In 2017, attorneys for Georgia-Pacific reorganized the company and redomiciled it in Texas so that it could utilize the part of the Texas Business Code that authorized firms to split through the Texas Divisive Merger Provisions. In doing so, the old Georgia-Pacific ("Old GP") reorganized into two firms, Bestwall and New Georgia-Pacific ("New GP"). Each of

these firms became wholly owned subsidiaries of Georgia-Pacific Holdings, LLC. Under Texas law, the assets and liabilities of the former entity could then be split between the newly created entities. As a result, all liability claims from Old GP's asbestos litigation were transferred to Bestwall along with a funding agreement between New GP and Bestwall. The funding agreement was Bestwall's only nontrivial asset and it included an arrangement to fund the asbestos litigation settlements. The issue herein is in the terms of the funding agreement between New GP (Payor) and Bestwall (Payee) relative to the realistic amount of damages that would result from the trend in verdicts against Old GP in cases that had been tried prior to the 2017 reorganization.

Case Study 2: Johnson & Johnson

Founded in 1886 after a inventing a breakthrough in surgical cotton, Johnson & Johnson (JNJ) quickly became a leader in industries that supplied surgeons during the scientific revolution of the late 19th Century. In addition to surgical cotton, JNJ also sold pain medications, bandages, needles in its inceptive years after being founded by three brothers—Robert Wood Johnson, James Wood and Edward Mead. The company grew rapidly to include other products that aided patients with the most innovative focused on the sterilization process to better help surgical wounds. JNJ became one of the first U.S. companies to codify a corporate code of conduct in 1943, the same year that it became publicly traded on the New York Stock Exchange (NYSE).[3] An early JNJ product was baby powder, invented in 1892 by the firm's scientific director Fred Kilmer (Oppenheimer 2013). Baby powder's main component is talc, which is a naturally occurring mineral composed of magnesium, silicone, oxygen and hydrogen (www.fda.gov). While natural, the use of this mineral has been controversial since at least the early 1970s when a possible link between the use of talc and ovarian cancer in woman was posited (Henderson, Joslin, Griffiths and

[3] Data collected from JNJ's website at www.ourstory.jnj.com

Turnbull 1971). By the early 1980s, numerous studies had found evidence for a statistically significant link between the consistent use of products containing talc and ovarian cancer (Cramer, Welch, Scully and Wojciechowski 1982).

From the 1970s through the early 2000s, an academic and regulatory tug-of-war took place where research studies would confirm or refute the early talc findings and regulatory agencies, such as the Food and Drug Administration (FDA), would publish ambiguous statements on the potential harm from the mineral. It was not until 2009 that the first talc lawsuit was initiated against JNJ by Deane Berg, who subsequently was awarded a verdict in her favor in 2013 but with no associated damages. Three years later, the first damages ($72 million) from a talc claim were rendered against the firm in a case filed by the Estate of Jacqueline Fox in Missouri's state court (NBC News, 16 February 2016). At the same time, hundreds of additional lawsuits were filed by plaintiffs across the United States, eventually growing to the tens of thousands in short order.

In October of 2021, JNJ announced on its website that it had formed a subsidiary named LTL Management LLC (LTL) that was formed for the sole purpose of holding the liability claims from the mass tort litigation due to talc (www.jnj.com) and that LTL was filing for Chapter 11 bankruptcy protection. At the time of the filing, JNJ faced approximately 40,000 legal claims due to talc and, as with Georgia-Pacific, it was attempting to shield the larger, publicly traded firm from bankruptcy jurisdiction through the Texas Two Step procedure. LTL was formed as a North Carolina corporation and filed for Chapter 11 in the U.S. Bankruptcy Court, Western District of North Carolina but was quickly removed to the U.S. Bankruptcy Court in New Jersey because, according to the Bankruptcy Administrator in North Carolina, jurisdiction in North Carolina was "…barely proper…" since LTL "…created facts to fit the statute (*In Re LTL Management LLC*: Motion of Bankruptcy Administrator To Transfer Venue, 10/25/21). On November 16, 2022, U.S.

Bankruptcy Judge J. Craig Whitley ordered the case transferred to New Jersey since most litigation had been filed in that state (*In Re LTL Management LLC*: Order 11/16/21). New Jersey was the domiciled location of JNJ, which was the named defendant in the thousands of talc lawsuits filed as of the bankruptcy date.

LTL was formed shortly before the bankruptcy filing as a Texas corporation. It was one of two entities formed in Texas spun off from Johnson & Johnson Consumer Inc (Old JJCI), a subsidiary of JNJ. In addition to LTL, a second entity was formed (New JJCI) and was immediately merged into JNJ. In essence, the former JJCI that was liable or potentially liable for the talc claims split itself into a new entity (LTL) whose only purpose was to hold the liability claims and then a second new entity (New JJCI) that held all assets and liabilities of Old JJCI except for the talc liability claims. LTL then moved its address to North Carolina and New JJCI, through its merger with the parent, was then domiciled in New Jersey (*In Re LTL Management LLC*: Order 11/16/21).

Case Study 3: 3M

3M is a U.S. materials company founded in 1902 as an anthracite miner. Originally called Minnesota Mining and Manufacturing Company, it became more widely known after the 1914 product release of Three-M-Ite Elek-Tro-Cut utility rag. 3M's notoriety then grew in 1925 when it created Scotch tape, a product it is still known for a century later. Following other large firms in the first part of the Twentieth Century such as Radio Corporation of American (RCA) and American Telephone & Telegraph (AT&T), 3M formed a research laboratory to test ideas for breakthrough technologies through extensive research and development. The firm was publicly listed in 1946 after success with some products used by American troops in World War II. Other key inventions and products included Fluorel, Thinsulate, Post-It Notes and Scotch-Brite all of

which made the company one of the world's largest with a peak market capitalization of $140 Billion (2018) and peak annual revenues of $35 Billion (2021).

In 2008, 3M acquired Aearo Technologies Inc., an Indianapolis-based protection equipment manufacturer (3M Press Release, 11/15/2007). Aearo Technologies was a leader in hearing and eyesight protective gear and supplied other manufacturers in the first responder, industrial and military market segments. Operating in 70 countries, Aearo had 1,700 employees at the time of the acquisition announcement and held patents on many of the products that were the basis of 3M's desire to acquire. One of the key market segments that Aearo was a leader in was that of energy absorption, especially in products related to hearing. After 3M absorbed Aearo, sales of its CAEv2 earplug to the military expanded with millions of units sold to multiple branches of the U.S. military.

3M's CAEv2 earplug was a two-sided earplug in which one side (yellow) was meant to restrict loud noises common in military situations, such as explosions, while allowing normal noises, such as verbal commands, to be heard. The other side of the earplug (olive) was intended to completely block noise completely. The differentiating aspect of the CAEv2 was in the yellow side, because it would presumably stop soldiers from hearing loss due to gunfire or explosions but allow regular communications, which are crucial for efficiency in the battlefield. However, many military members began to complain of hearing loss after using the product, leading to mass tort litigation in the U.S. federal court system. In 2021, a series of initial (aka bellweather) cases were decided. Three of the five of the lawsuits ended in a plaintiff's victory against 3M with total damages amounting to $17 million, while two of the actions ended in a 3M victory.

Similar to Georgia-Pacific and Johnson & Johnson, 3M decided to initiate a bankruptcy filing similar to the Texas Two Step strategy to combat mass tort litigation, which had ballooned

to nearly 300,000 claims by 2022. One difference between 3M and the other two subjects of the preceding cases was that Aearo Technologies was an existing subsidiary of 3M as opposed to a newly created subsidiary. However, a major similarity to the Georgia-Pacific and Johnson & Johnson narratives was that the large, publicly traded 3M attempted to circumvent the bankruptcy process itself even though it was a defendant in the mass tort liability lawsuits filed through the date of Aearo's bankruptcy petition.

DISCUSSION

Findings

Applying the theoretical concepts in the current work to the qualitative empirical evidence also presented herein is crucial to the study of objective ethics. As proposed, firms may violate ethical norms through hypocrisy if they fail to adhere to the notion of action-word reconciliation and/or they are the proximate cause of negative externalities. In this section, each of the firm's actions regarding the mass tort litigation and subsequent bankruptcy filing will be analyzed further in the context of each of the subject firm's codified principles. At a minimum, each firm's Code of Conduct (COC) for the two years prior to the bankruptcy filing were analyzed. Other sources of proactive corporate statements were also researched including press releases, annual proxy statements and Corporate Social Responsibility (CSR) reports.

REFERENCES

3M News Release. 3M To Acquire Aerao Technologies Inc. 3M Website 15 November 2007.

Arli, D., Grace, A., Palmer, J. and Pham, C. 2017. Investigating the Direct and Indirect Effects of Corporate Hypocrisy and Perceived Corporate Reputation On Consumers' Attitudes Towards the Company. *Journal of Retailing and Consumer Services* 37: 139-45.

Author Unknown. 1975. Johns-Manville Cited Again In Class Action Over Asbestos Hazards. *Wall Street Journal* 27 August 1975.

Biglan, A. 2009. "The Role of Advocacy Organizations in Reducing Negative Externalities." *Journal of Organizational Behavior Management*, 29(3): 215-230.

Brown, R. 2018. Congressional Certification: How Firms Use Former Congressional Members To To Lobby On Their Behalf. *Journal of Public Affairs*, 18.

Brown, R. 2022. Cronyism As Strategy: How Firms Co-opt Government and Privatize Rents. *Journal of Managerial Issues*, Forthcoming.

Chang, D. 2016. The Sticky Side of Hypocrisy: Environmental Activists in an Oil Drenched World. *Philosophical Inquiry In Education*, 23(2): 200-202.

Chen, C. and X. Chen. 2009. "Negative Externalities of Close Guanxi Within Organizations." *Asia Pacific Journal of Management*, 26: 37-53.

Cook, W. 1924. Fibrosis of the Lungs Due To The Inhalation of Asbestos Dust. *British Medical Journal*, 2: 147.

Cramer, D., Welch, W., Scully, R. and Wojciechowski, C. 1982. Ovarian Cancer and Talc: A Case Control Study. *Cancer*, 50(2): 372-76.

Demir, E., Y. Huang, S. Scholts, and T. Van Woensel. 2015. "A Selected Review on the

Negative Externalities of the Freight Transportation: Modeling and Pricing." *Transportation Research Part E*, 77: 95-114.

Diao, M., Y. Qin, and T. Sing. 2016. "Negative Externalities of Rail Noise and Housing Values: Evidence From the Cessation of Railway Operations in Singapore." *Real Estate Economics*, 44(4): 878-917.

Donegan, J., M. Ganon, and Z. Johnson. 2017. "Influencers of Earnings Management Fraud: Coercion, Project Stage and Ethics." *Journal of Managerial Issues*, 29(2): 169-188.

Dubbink, W. and Van Liedekerke, L. 2020. Rethinking The Purity of Moral Motives In Business: Kant Against Moral Purism. *Journal of Business Ethics* 167: 379-393.

Duke, J. and D. Sassoon. 2017. "A Classroom Game on a Negative Externality Correcting Tax: Revenue Return, Regressivity and the Double Dividend." *The Journal of Economic Education*, 48(2): 65-73.

Enderwick, P. 2005. "What's bad about crony capitalism?" *Asian Business & Management*, 4: 117-132.

Epstein, R. 2018. "Positive and Negative Externalities in Real Estate Development." *Minnesota Law Review*, 102: 1493-1525.

Freyman, N. 2021. Hertz Shareholders Win Big With Bankruptcy Deal. *Morning Brew* 21 May 2021.

Friedland, J. 2005. Confirming a Plan. *American Bankruptcy Institute Journal*, (23):10.

Griffiths, M., B. Perera. and P. Albinsson. 2018. "Contrived Surplus and Negative Externalities in the Sharing Economy." *Journal of Marketing Theory and Practice*, 27(4): 445-463.

Henderson, W., Joslin, C., Griffiths, K. and Turnbull, A. 1971. Talc and Carcinoma of the Ovary and Cervix. *BJOG: An International Journal of Obstetrics and Gynaecology*, 78(3): 266-

272.

Jauernig, J., Uhl, M. and Valentinov, V. 2021. The Ethics of Corporate Hypocrisy: An Experimental Approach. *Futures*, 131.

Kessler, R. 1968. Medical Researchers Study Role of Asbestos As Cause of Ailments. *Wall Street Journal* 7 November 1968.

Krekel, C. and A. Zerrahn, 2017. "Does the Presence of Wind Turbines Have Negative Externalities For People in Their Surroundings?" Evidence From Well-Being Data. *Journal of Environmental Economics and Management*, 82: 221-238.

Lenz, H. 2020. Aggressive Tax Avoidance By Managers of Multinational Companies As a Violation of Their Moral Duty To Obey the Law: A Kantian Rationale. *Journal of Business Ethics* 165: 681-697.

Losado-Otalora, M. and Alkire, L. 2021. A Transformative Approach To Corporate Social Responsibility: An Antidote To Corporate Hypocrisy. *The Service Industries Journal*, 41(3/4): 200-222.

Manish, G. P. and D. Sutter. 2016. "Mastery Versus Profit As Motivation For the Entrepreneur: How Crony Policies Shape Business." *Journal of Entrepreneurship and Public Policy*, 5: 95-112.

Mansell, S. 2013. Shareholder Theory and Kant's Duty of Beneficence. *Journal of Business Ethics* 117: 583-599.

Nagler, M. 2011. "Negative Externalities, Competition and Consumer Choice." *The Journal of Industrial Economics*, 59(3): 396-421.

NBC News. 2016. Court Orders Johnson & Johnson To Pay $72M In Talcum Powder Ovarian Cancer Case. From www.nbcnews.com 16 February 2016.

O'Donnell, M. 1976. Henry Sidgwick: Externalities and the Emergence of the New Concept of Laissez-Faire. Doctoral Dissertation, Texas A&M University.

Official Committee of Asbestos Claimants and The Future Claimants' Representative v. Bestwall LLC et.al. United States Bankruptcy Court, Western District of North Carolina. 29 July 2020.

Oppenheimer, J. 2013. Crazy Rich: Power, Scandal and Tragedy Inside the Johnson & Johnson Dynasty. St. Martin's Press New York, NY.

Plevin, M. et.al. 2016. Where Are They Now, Part Eight: An Update On Developments In Asbestos-Related Bankruptcy Cases. *Mealy's Asbestos Bankruptcy Report* 16(2): 1-57.

Scheidler, S., Edinger-Schons, L., Spanjol, J. and Wieseke, J. 2019. Scrooge Posing As Mother Teresa: How Hypocritical Social Responsibility Strategies Hurt Employees and Firms. *Journal of Business Ethics* 157: 339-358.

Scotti, D., M. Dresner, G. Martini, and C. Yu. 2014. "Incorporating Negative Externalities Into Productivity Assessments of U.S. Airports." *Transportation Research Part A*, 62: 39-53.

Wagner, T., Lutz, R. and Weitz, B. 2009. Corporate Hypocrisy: Overcoming the Threat Of Inconsistent Corporate Social Responsibility Perceptions. *Journal of Marketing* 73(6): 77-91.

Wootton, D. 2018. Power, Pleasure, and Profit. Belknap/Harvard Publishing.

Yang, L. 2019. "Negative Externality of Fiscal Problems: Dissecting the Contagion Effect of Municipal Bankruptcy." *Public Administration Review*, 79(2): 156-167.

Websites Cited

3M

Georgia-Pacific

Johnson & Johnson, www.ourstory.jnj.com

Texas State Historical Association, https://www.tshaonline.org/handbook/entries/borel-v-fibreboard-paper-products-corporation

Operations Management / Operations Research

AN EOQ MODEL WITH PLANNED SHORTAGES AND GEOMETRICALLY DECLINING VARIABLE PROCUREMENT COST PER UNIT

ABSTRACT

This paper extends the seminal Economic Order Quantity model that allows backlogging of out-of-stock units by assuming that the unit variable procurement cost for each item in the lot is not fixed, but rather declines geometrically as the lot size increases. Based on this scenario, we present a modified EOQ model with planned shortages and provide explicit expressions for the approximately optimal values of decision variables.

Keywords: EOQ with Planned Shortage Model, Procurement Cost

INTRODUCTION

The classical EOQ model, frequently referred to as Wilson's Square Root Model, is based on a series of highly restrictive assumptions among which are constant deterministic demand, constant holding costs across time, perfect quality of replenishment items, and constant procurement (manufacturing) cost. These assumptions limit the applicability of the Wilson's model to the actual inventory systems that one encounters in organizations. In recognition of this fact, a vast number of researchers in the area of inventory management have developed more realistic inventory models by relaxing these assumptions in various ways. In this current work, we focus on the constant procurement cost function assumption. In Wilson's model, this assumption results in the procurement cost disappearing from the model as a result of the process of classical optimization. However, it is clearly possible that this cost may not be constant. For example, process quality may be improved through the life of the cycle by elimination of special causes, operator performance may be improved by training and

learning effects, machine maintenance may improve performance, etc. Under these scenarios it is reasonable to assume that the cost of procurement will decrease for each additional unit procured (manufactured). This research begins the investigation of variable procurement cost by proposing a model in which this cost decreases geometrically with increasing lot size.

MODEL AND ASSUMPTIONS

Consider the basic Economic Order Quantity model that allows backordering of out-of-stock units with the following total cycle cost and total annual cost functions, $C_{Basic}(S,Q)$ and $A_{Basic}(S,Q)$, respectively.

$$C_{Basic}(S,Q) = K + CQ + C_h \frac{(Q-S)^2}{2D} + \frac{S^2}{2D} C_b \quad (1)$$

$$A_{Basic}(S,Q) = \frac{D}{Q}K + CD + C_h \frac{(Q-S)^2}{2Q} + \frac{S^2}{2Q} C_b \quad (2)$$

Where

D = Annual Demand in units,
Q = Lot size per setup (order),
T = D/Q = Cycle time, the time between placement (or receipt) of replenishment orders,
S = Number of units backordered,
K = Setup (order) cost per setup (order),
C = Unit variable cost of procurement (production or purchase),
C_h = Inventory holding cost per unit per year, independent of dollar value,
C_b= Backordering cost per unit short per year.

The results of classical optimization yields the well-known expressions for the optimal values for the lot size, Q^*_{Basic}, units backordered, S^*_{Basic}, and the total annual cost, $A^*_{Basic}(S,Q)$ as follows:

$$Q^*_{Basic} = \sqrt{2DK\left(\frac{1}{C_h}+\frac{1}{C_b}\right)} \tag{3}$$

$$S^*_{Basic} = Q^*_{Basic}\left(\frac{C_h}{C_h + C_b}\right) \tag{4}$$

and

$$A^*_{Basic}(S,Q) = CD + \sqrt{\frac{2DK}{\frac{1}{C_h}+\frac{1}{C_b}}} \tag{5}$$

Implicit in these derivations is that the unit variable cost of procurement, C, is assumed to be the same (i.e. fixed or constant), for all units in the lot. Now, assume that this is not the case. Specifically, assume that the initial unit variable procurement cost for the first item in the lot, C_1, equals C. However, for reasons such as the effects of learning in production or the incremental unit price discounts in purchasing, the unit procurement cost of each of the subsequent items (C_j for j=2, 3, 4, …, Q) declines by a factor of r, $0 < r < 1$, compared to the unit procurement cost of the previous item (C_{j-1} for j= 2, 3, 4, …, Q). With these specifications, the unit variable cost of procurement for the j^{th} unit in each lot can be expressed as $C_j = Cr^{j-1}$ *for* $j = 1,2,3,...,Q$.

Based on the above scenario, we now modify the EOQ with planned shortages model based on the assumption that unit variable cost of procurement cost declines by a factor of r, $0 < r < 1$, as follows.

The total procurement (production or purchase) cost per cycle for Q units is K + V(Q), where V(Q) is the cumulative variable cost of procurement for a lot of size Q and is given by

$$V(Q)=\sum_{j=1}^{Q} Cr^{j-1}=C+Cr+Cr^2+Cr^3+...+Cr^{Q-1} \tag{6}$$

Please note that if r =1 in equation (6), then $V(Q)=CQ$, and the results of the basic model represented by equations (1) through (5) apply. For 0 < r < 1, equation (6) is the sum of the first Q terms of a geometric series, up to and including the Cr^{Q-1} term, with common ratio of any term with the previous one, r, and the coefficient, C, representing the first term of the geometric series in expanded form. Therefore, upon using the closed-form formula for the sum of the first Q terms, we find the cumulative variable cost of procurement per cycle as follows

$$V(Q)=\sum_{j=1}^{Q} Cr^{j-1}=\sum_{k=0}^{Q-1} Cr^{k}=C\left(\frac{1-r^{Q}}{1-r}\right) \tag{7}$$

Using (7), the total cost per cycle for the modified EOQ model with planned shortages and declining unit procurement cost is

$$\begin{aligned} C_{Modified}(S,Q) &= K+V(Q)+C_h\frac{(Q-S)^2}{2D}+\frac{S^2}{2D}C_b \\ &= K+C\left(\frac{1-r^{Q}}{1-r}\right)+C_h\frac{(Q-S)^2}{2D}+\frac{S^2}{2D}C_b \end{aligned} \tag{8}$$

Multiplying equation (8) by the number of cycles per year, D/Q, after some simplifications, we find the total annual cost for the modified model

$$A_{Modified}(S,Q)=\frac{D}{Q}K+\left(\frac{CD}{Q}\right)\left(\frac{1-r^{Q}}{1-r}\right)+C_h\frac{(Q-S)^2}{2Q}+\frac{S^2}{2Q}C_b \tag{9}$$

Typically, the rate of decline (1-r) in variable procurement cost per unit from one item in the lot to the next (from C_j to C_{j+1}, where $j=1,2,3,...,Q$) is low. Hence, the common ratio, r, is close to one. Under this condition, Porteus [1986], in the context of quality in EOQ model, indicates that $r^{Q}=e^{(\ln r)Q}$. This relationship was also used by Nasri, Paknejad, and Affisco [2006] in the context of quality as it pertains to a continuous review (s, Q) model. Using a second order Taylor series expansion of $r^{Q}=e^{(\ln r)Q}$, and

setting $\ln r \cong -\left(\frac{1-r}{r}\right)$, yields

$$r^{Q} = e^{(\ln r)Q} \cong 1-\left(\frac{1-r}{r}\right)Q+\left(\frac{1}{2}\right)\left[\left(\frac{1-r}{r}\right)Q\right]^{2} \tag{10}$$

Upon using (10) in (9) and simplifying, we find an approximate expression for the total annual cost of the modified model as follows:

$$A_{Modified}(S,Q) = \frac{D}{Q}K + \left(\frac{CD}{r}\right)\left[1-\left(\frac{1-r}{r}\right)\left(\frac{Q}{2}\right)\right] + C_h\frac{(Q-S)^2}{2Q}+\frac{S^2}{2Q}C_b \tag{11}$$

A necessary condition for $A_{Modified}(S,Q)$ in equation (11) to have a minima is that

$$\frac{\partial A_{Modified}(S,Q)}{\partial Q} = -\frac{DK}{Q^2}-\frac{CD}{2r}\left(\frac{1-r}{r}\right)+\frac{C_h}{2}-\frac{1}{2}\left(C_h+C_b\right)\left(\frac{S^2}{Q^2}\right)=0 \tag{12}$$

and

$$\frac{\partial A_{Modified}(S,Q)}{\partial S} = -C_h+\left(C_h+C_b\right)\left(\frac{S}{Q}\right)=0 \tag{13}$$

Solving (12) and (13) simultaneously, produces

$$Q^{*}_{Modified} = \sqrt{2DK\left(\frac{1}{C_h}+\frac{1}{C_b}\right)\left(\frac{1}{1-\alpha}\right)} \tag{14}$$

and

$$S^{*}_{Modified} = Q^{*}_{Modified}\left(\frac{C_h}{C_h+C_b}\right) \tag{15}$$

Where $0 \le \alpha < 1$ and is given by

$$\alpha = CD\left(\frac{1-r}{r^2}\right)\left(\frac{1}{C_h}+\frac{1}{C_b}\right) \tag{16}$$

Please note that if $r=1$, then, $\alpha = 0$, and (14) and (15) become identical to the corresponding results of the basic EOQ model with planned shortages given by equations

(3) and (4). Finally, to show that the stationary point $\left(Q^*_{Modified}, S^*_{Modified}\right)$ is minimum, it is sufficient to show that the first and second principal minors of the Hessian determinant of $A_{\mathrm{mod}ified}(S,Q)$ function given in equation (11) are both strictly positive. The first principle minor is

$$\left|H_{11}\right| = \frac{\partial^2 A_{\mathrm{mod}ified}}{\partial Q^2} = \frac{2DK}{Q^3} + \left(C_h + C_b\right)\left(\frac{S^2}{Q^3}\right) \tag{17}$$

Which is, of course, strictly positive at $\left(Q^*_{Modified}, S^*_{Modified}\right)$ when $0 \le \alpha < 1$.

The second principle minor is

$$\left|H_{22}\right| = \begin{vmatrix} \dfrac{\partial^2 A_{\mathrm{mod}ified}}{\partial Q^2} & \dfrac{\partial^2 A_{\mathrm{mod}ified}}{\partial Q \partial S} \\ \dfrac{\partial^2 A_{\mathrm{mod}ified}}{\partial Q \partial S} & \dfrac{\partial^2 A_{\mathrm{mod}ified}}{\partial S^2} \end{vmatrix} = \frac{2DK}{Q^4}\left(C_h + C_b\right) \tag{18}$$

Which is again strictly positive at the stationary point $\left(Q^*_{Modified}, S^*_{Modified}\right)$ given by equations (14) and (15).

CONCLUSION

This paper opens a new line of inquiry into order quantity models under the situation of declining variable cost of procurement (production or purchase). Approximately optimal expressions, in closed forms, are derived for policy variables of Economic Order Quantity model with planned shortages and geometrically declining unit variable procurement cost. It is shown that the traditional EOQ model with planned shortages and constant unit variable procurement cost can be treated as a special case of the results presented in this paper.

Acknowledgement: This research was partially supported by a Special Leave Grant from the Frank G. Zarb School of Business, Hofstra University.

REFERENCES

Nasri, F., Paknejad, J., & Affisco, J.F. (2006). An Approximately Optimal Quality Adjusted Continuous Review (s,Q) Inventory Model. *Northeast Decision Sciences Institute Proceedings*, pp. 469-476.

Porteus, E. (1986) “Optimal lot sizing, process quality improvement and setup cost Reduction. *Operations Research*. 34(1), 137-144.

Competitive Balance in Major League Baseball and the National Basketball Association: A Comparison of the Two Leagues

Abstract

Professional sport leagues, unlike other business enterprises, depend on stiff competition for economic survival. Sports are entertainment; tight division races, unpredictable playoff series, and the periodic emergence of new champions enhance the entertainment value of the sport, ensuring the league's future fan base. However, while individual teams need the league to succeed, winning is the key to their economic success. Winning increases fan interest, brings more people to the stadium or arena, improves television ratings, and bolsters sales of team-related merchandise, all of which add to the team's prosperity.

Major League Baseball (MLB) entered the era of free agency on December 23, 1975, and player salaries have since grown to extraordinary levels. In the summer of 1988, the National Basketball Association (NBA) collective bargaining agreement introduced the concept of unrestricted free agency, which has led to similar average player salary growth. Some teams in MLB and in the NBA, notably those located in larger markets and those possessing greater financial resources, found it easier than other teams to sign free agents to high-salary, multi-year contracts, thereby cornering the market on the most talented players and threatening the competitive balance on the field or court.

In this paper, we measure MLB and NBA front office, on-field (or on-court), and overall team performance over the past two decades and use these measures to determine if there is competitive balance in the leagues.

1. Introduction

Professional sport leagues, unlike other business enterprises, depend on stiff competition for economic survival. Sports are entertainment; tight division races, unpredictable playoff series, and the periodic emergence of new champions enhance the entertainment value of the sport, ensuring the league's future fan base. However, while individual teams need the league to succeed, winning is the key to their economic success. Winning increases fan interest, brings more people to the stadium or arena, improves television ratings, and bolsters sales of team-related merchandise, all of which add to the team's prosperity.

Major League Baseball (MLB) entered the era of free agency on December 23, 1975, and player salaries have since grown to extraordinary levels. In 1975, the average player salary in MLB was $44,676; in 2018, it was more than 100 times greater at $4.52 million. In the summer of 1988, the National Basketball Association (NBA) collective bargaining agreement introduced the concept of unrestricted free agency, which has led to similar average player salary growth. In 1988, the average player salary in the NBA was $528,011; in 2018, it was $7.42 million. Some teams in MLB and in the NBA, notably those located in larger markets and those possessing greater financial resources, found it easier than other teams to sign free agents to high-salary, multi-year contracts, thereby cornering the market on the most talented players and threatening the competitive balance on the field or court.

In response, MLB instituted its first luxury tax as part of the 1996 collective bargaining agreement. In July 2000, the *Commissioner's Blue Ribbon Panel on Baseball Economics* (Levin et al., 2000) reported on the revenue disparities in MLB. The Panel found that these disparities were affecting competition, that the disparities were becoming worse, and that the limited revenue sharing and payroll taxes approved in the 1996 labor agreement with the players were having little effect. Moreover, the Panel concluded that the cost of trying to be competitive was raising ticket and concession prices, jeopardizing MLB's position as the affordable family spectator sport. The Panel's recommendations included greater revenue sharing and a competitive balance tax, both of which were part of the 2002 labor agreement with the players.

The NBA initiated a salary cap on team payrolls in 1984. Unlike the "hard" cap in the National Football League and the National Hockey League, the NBA has a "soft" cap that permits teams to exceed the cap under certain circumstances. In addition, the NBA imposes a luxury tax on teams whose payroll exceeds the salary cap by a specified amount.

The 2018 opening day total player salary for the San Francisco Giants was $221,426,944 while that of the Oakland Athletics was $62,652,500. With one team's total player salary equal to 3.53 times that of another team (in essentially the same market), it is reasonable to ask whether the team with the lower salary can effectively compete with the team with the higher salary, and the extent to which market size influences competitiveness. In this paper, we measure MLB and NBA front office, on-field (or on-court), and overall team performance over the past two decades and use these measures to determine if there is competitive balance in the leagues. Specifically, we pose the following research questions for each season in the study period:

1. How many MLB and NBA teams are noncompetitive due to low total player salary?
2. How many MLB and NBA teams are overspending on total player salary?
3. How much does an MLB and an NBA team need to spend on total player salary to be competitive?
4. What is the maximum total player salary that an MLB and an NBA team can pay without overspending?
5. How does noncompetitiveness due to low total player salary relate to market size in MLB and in the NBA?
6. How does overspending on total player salary relate to market size in MLB and in the NBA?
7. Is there competitive balance in MLB and in the NBA?

2. Literature Review

Data envelopment analysis (DEA) has been extensively used in operations research and economics to measure the productive efficiency of *decision making units* (DMUs.) It has been implemented in a variety of sports such as baseball, basketball, football, soccer, hockey, golf, tennis, and the summer and winter Olympics in order to quantify the performance of individual athletes, teams, and countries. The most common DEA models are single stage models that measure the efficiency with which an organization converts multiple inputs into multiple outputs. However, in order to measure team efficiency in Major League Baseball (MLB) and in the National Basketball Association (NBA) we employ a multi-stage network DEA model. This model will enable us to analyze both the front office and on-field (on-court) management of each team to identify areas of strength and weakness.

Leach and Barrows (2007) used DEA to analyze the impact of management in the English Premier League Football clubs. To do this they used both on field variables such as assists and goals scored and managerial variables such as the financials of the teams to form a frontier for best managerial practices

and see how all the teams compared to this frontier. Gutierrez and Ruiz (2013) used DEA to assess the individual offensive performances of 66 players in the Spanish Premier League. This analysis allowed them to rate the relative strengths and weaknesses of the given players and to derive a full ranking. Lozano et. All (2002) used DEA to measure the performance of nations at the past five summer Olympics. To do this they used two inputs, GNP and population and three outputs, the number of the three different medals. Higher weights were given to the higher ranking medals for the sake of consistency. VRS was assumed and the results were plotted for each country. This showed their relative strengths and weaknesses.

There have been various studies that have had the similar focus of studying the efficiency of NBA and MLB franchises. Einolf (2004) used DEA to measure franchise payroll efficiency in the National Football League (NFL) from 1981 to 2000 and Major League Baseball (MLB) from 1985 to 2001. Results showed that because of the implementation of the salary cap, NFL franchise payrolls became more efficient after its inception while MLB franchise payrolls remained less efficient. Porter and Scully (1982) use DEA to estimate the managerial efficiency in MLB. They used slugging averages and pitcher strikeout ratios as inputs and wins as an output to determine the best overall manager for the study period.

Anderson and Sharpe (1997) used DEA to create an alternative to traditional batting statistics called the Composite Batter Index (CBI). This had various advantages over traditional statistics since it was a relative measure that compared players. They also uses historical results to examine how batting has changed over the years. Ruggiero (2011) used DEA to develop the theory of DEA in the context of a production environment. The main focus was to assess both the technical and cost efficiency of MLB teams. The same model was used to analyze Hall of Famers and see how deserving they are of the honor.

Cooper, Ruiz, and Sirvent (2009) used DEA to identify component profiles as well as overall indices of performance in the context of an application to assessments of basketball players. Their focus was on the multiplier values for the efficiently rated players. For this, they use a procedure that they recently developed that guarantees a full profile of non-zero weights, or "multipliers." They demonstrate how these values can be used to identify relative strengths and weaknesses in individual players.

Savic, Radojicic, and Radovanovic (2014) used DEA to evaluate an efficiency of the NBA players, based on multiple inputs and multiple outputs. The efficiency is evaluated for 26 NBA players at the guard position based on existing data. Since the model used existing data, it would not be applicable to a new player. To accomplish this, they used machine learning algorithms to predict the efficiency of any new players. Bolen, Rezek, and Pitts (2017) used DEA to examine the efficiency of men’s college basketball

programs in producing competitive and academically successful teams. They used inputs measuring team talent, academic ability, and experience. Two efficiency measures are estimated which rank teams according to their technical efficiency in producing each of the outputs desired by athletic departments and universities.

There are four studies using DEA that are relevant to this paper. Lewis and Sexton (2004) use DEA to analyze all 30 MLB organizations during the 1999 regular season to determine their on-field and front office relative efficiencies. Their on-field DEA model employs one output (wins) and two symmetrically defined inputs, one to measure offense and one to measure defense. The defensive measure is such that larger values correspond to worse defensive performance, rather than better, and hence is a reverse input. The front office model uses one input (team salary) and two outputs, one is a reverse output, which are the inputs to the on-field model. The next two papers are from Lewis, Lock, and Sexton (2007) and (2009). In Lewis, Lock, and Sexton (2007), a two stage DEA model is used to determine the minimum player salary required for a team to be competitive in every non-strike season s in the study period. They use this to determine which teams are competitive each season and the relationship between competitiveness and market size. In Lewis, Lock, and Sexton (2009), Network DEA is used to derive managerial efficiency scores for all MLB teams from 1901 to 2002. Team capability is measured by analyzing offensive and defensive statistics. They define each team's effectiveness by its win percentage. They discover that capability has a bigger impact on a team's regular season effectiveness. The last paper is Moreno and Lozano (2012) that uses a Network-DEA approach to consider the distribution of the budget between first-team players and the rest of the payroll. The approach consists of five stages, which evaluate the performance of first-team and bench-team players, the offensive and defensive systems and the ability for transforming the points made by the team and by the opponents into wins. The model applied to the 30 NBA teams for the regular season 2009–2010. The results show that network DEA has more discriminating power and provides more insight than the conventional DEA approach.

In this paper, we replicate the analysis in Lewis, Lock, and Sexton (2007) on a slightly more complex network model representation of an MLB team, for the seasons 2001 to 2018. We develop an analogous network model for an NBA team and apply the methodology for the seasons 2004-05 to 2017-18. Finally, we compare competitive balance in the two leagues.

3. Methodology

The Commissioner's Blue Ribbon Panel (Levin et al., 2000) defines competitive balance as the state in which *"… every well-run club has a regularly recurring reasonable hope of reaching post-season play."* Our analysis entails parsing this statement into operational definitions of *"well-run"* and *"reasonable hope of reaching post-season play."*

3.1 Efficiency Measurement

To define *"well-run,"* we turn to the theory of productive efficiency in the management science and economics literature. We apply Network DEA to compute the efficiency of every MLB and NBA team in the study period relative to the frontier created by all other teams in the same season. The multi-stage production model is presented in Figure 1.

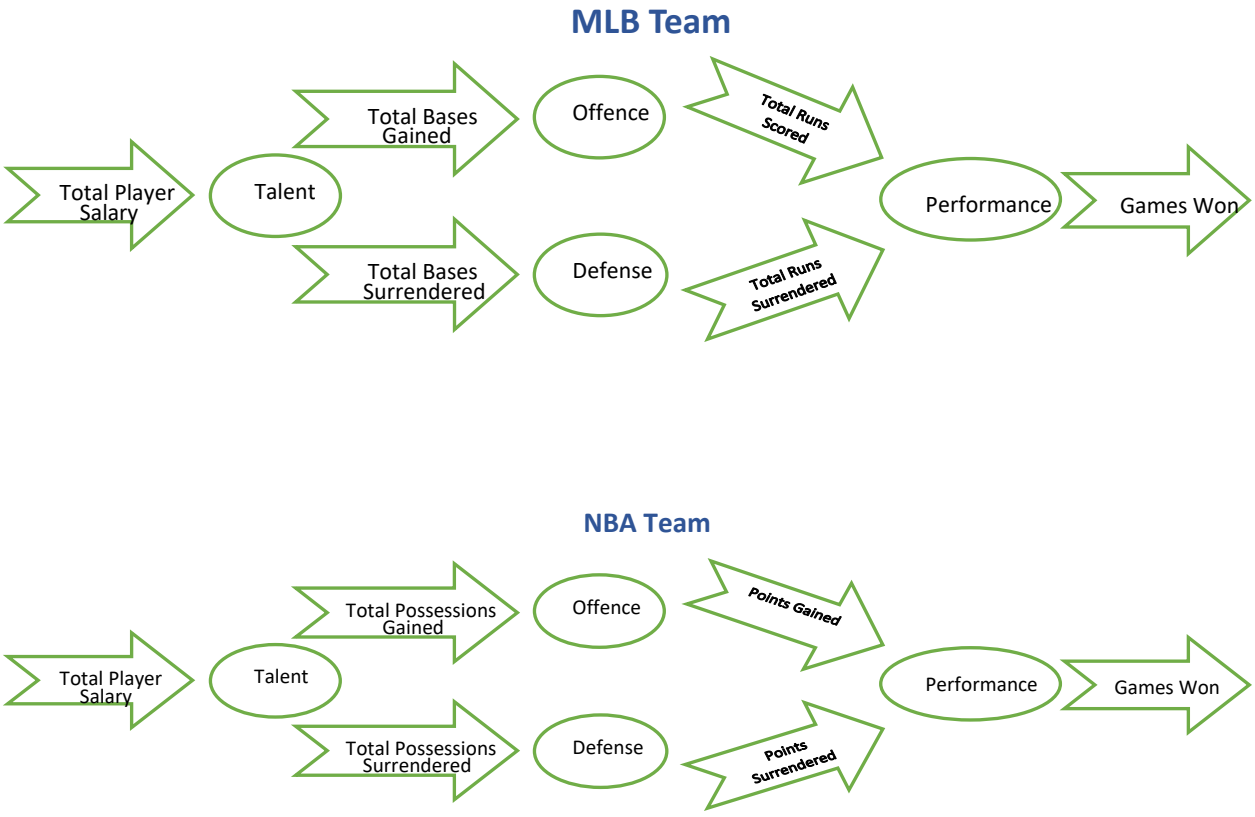


Figure 3.1: Sequential multi-stage model of an MLB and NBA team consisting of a front office operation and an on-field (on-court) operation.

3.1.1 Definitions

Let SAL_j denote the *total player salary* of team j in a given season and GW_j demote the *games won* by team j in a given season. Let TBG_j denote the *total bases gained* by MLB team j in a given season, which is a combination of a team's singles, doubles, triples, home runs, walks, and errors forced and TBS_j denote the *total bases surrendered* by MLB team j in a given season, which is a combination of opponents' singles, doubles, triples, home runs, walks, and the team's errors committed. Let RG_j denote the *runs gained* by MLB team j in a given season and RS_j denote the *runs surrendered* by MLB team j in a given season. Let TPG_j denote the *total possessions gained* by NBA team j in a given season, which is a combination of a team's offensive rebounds, defensive rebounds, steals, and turnovers forced and TPS_j denote the *total possessions surrendered* by NBA team j in a given season, which is a combination of opponents' offensive and defensive rebounds, steals, and the team's turnovers. Let PG_j denote the *points gained* by NBA team j in a given season including free throws, 2-point field goals, and 3-pointers made and PS_j denote the *points surrendered* by NBA team j in a given season including free throws, 2-point field goals, and 3-pointers given up.

Define θ_{1k} to be the *inverse efficiency of the front office* for team k, ε_{1k} to be the *efficiency of the front office* for team k, θ_{2k} to be the *inverse efficiency of the on-field (on-court) offense* for team k, ε_{3k} to be the *efficiency of the on-field (on-court) defense* for team k, and θ_{4k} to be the *inverse efficiency of the on-field (on-court) integration* for team k. Further, define λ_{1j} to be the *weight placed on the front office* of team j by team k, λ_{2j} to be the *weight placed on the on-field (on-court) offense* of team j by team k, λ_{3j} to be the *weight placed on the on-field (on-court) defense* of team j by team k, and λ_{4j} to be the *weight placed on the on-field (on-court) integration* of team j by team k.

3.1.2 MLB Network DEA Model

First, we solve the front office DEA model for MLB team k.

$$\text{Maximize } \theta_{1k}$$

$$\text{subject to}$$

$$\sum_{j=1}^{n} \lambda_{1j} SAL_j \leq SAL_k$$

$$\sum_{j=1}^{n} \lambda_{1j} TBG_j \geq \theta_{1k} TBG_k$$

$$\sum_{j=1}^{n} \lambda_{1j} TBS_j \leq \epsilon_{1k} TBS_k$$

$$\sum_{j=1}^{n} \lambda_{1j} = 1$$

$$\theta_{1k} + \epsilon_{1k} = 2$$

$$\lambda_{1j} \geq 0; \;\; j = 1,2, \dots, n$$

$$\theta_{1k}, \epsilon_{1k} \geq 0$$

Next, we solve the on-field offense and on-field defense DEA models for MLB team k.

Maximize θ_{2k}

subject to

$$\sum_{j=1}^{n} \lambda_{2j} TBG_j \leq TBG_k$$

$$\sum_{j=1}^{n} \lambda_{2j} RG_j \geq \theta_{2k} RG_k$$

$$\sum_{j=1}^{n} \lambda_{2j} = 1$$

$$\lambda_{2j} \geq 0; \;\; j = 1,2, \dots, n$$

$$\theta_{2k} \geq 0$$

Minimize ε_{3k}

subject to

$$\sum_{j=1}^{n} \lambda_{3j} TBS_j \geq TBS_k$$

$$\sum_{j=1}^{n} \lambda_{3j} RS_j \leq \varepsilon_{3k} RS_k$$

$$\sum_{j=1}^{n} \lambda_{3j} = 1$$

$$\lambda_{3j} \geq 0; \;\; j = 1,2, \ldots, n$$
$$\varepsilon_{3k} \geq 0$$

Finally, we solve the on-field integration DEA model for MLB team k.

Maximize θ_{4k}

subject to

$$\sum_{j=1}^{n} \lambda_{4j} RG_j \leq RG_k$$
$$\sum_{j=1}^{n} \lambda_{4j} RS_j \geq RS_k$$
$$\sum_{j=1}^{n} \lambda_{4j} GW_j \geq \theta_{4k} GW_k$$
$$\sum_{j=1}^{n} \lambda_{4j} = 1$$
$$\lambda_{4j} \geq 0; \;\; j = 1,2, \ldots, n$$
$$\theta_{4k} \geq 0$$

Let $TBG_k^* = \sum_{j=1}^{n} \lambda_{1j}^* TBS_j$ and $TBS_k^* = \sum_{j=1}^{n} \lambda_{1j}^* TBS_j$ where λ_{1j}^* is the optimal weight obtained when solving the front office model for team k. Now, we resolve the on-field offense model for team k using TBG_k^* as the RHS of the first constraint and resolve the on-field defense model for team k using TBS_k^* as the RHS of the first constraint.

Maximize θ_{2k}

subject to

$$\sum_{j=1}^{n} \lambda_{2j} TBG_j \leq TBG_k^*$$
$$\sum_{j=1}^{n} \lambda_{2j} RG_j \geq \theta_{2k} RG_k$$
$$\sum_{j=1}^{n} \lambda_{2j} = 1$$
$$\lambda_{2j} \geq 0; \;\; j = 1,2, \ldots, n$$
$$\theta_{2k} \geq 0$$

Minimize ε_{3k}

subject to

$$\sum_{j=1}^{n} \lambda_{3j} TBS_j \geq TBS_k^*$$

$$\sum_{j=1}^{n} \lambda_{3j} RS_j \leq \varepsilon_{3k} RS_k$$

$$\sum_{j=1}^{n} \lambda_{3j} = 1$$

$$\lambda_{3j} \geq 0; \;\; j = 1,2, \ldots, n$$

$$\varepsilon_{3k} \geq 0$$

Let ${}^*RG_k^* = \sum_{j=1}^{n} {}^*\lambda_{2j}^* RG_j$ where ${}^*\lambda_{2j}^*$ are the optimal weights obtained when solving the on-field offense model for team k, assuming the front office is efficient and ${}^*RS_k^* = \sum_{j=1}^{n} {}^*\lambda_{3j}^* RS_j$ where ${}^*\lambda_{3j}^*$ are the optimal weights obtained when solving the on-field defense model for team k, assuming the front office is efficient. Finally, we resolve the on-field integration model for team k using ${}^*RG_k^*$ and ${}^*RS_k^*$ as the RHS of the first and second constraints, respectively.

Maximize θ_{4k}

subject to

$$\sum_{j=1}^{n} \lambda_{4j} RG_j \leq {}^*RG_k^*$$

$$\sum_{j=1}^{n} \lambda_{4j} RS_j \geq {}^*RS_k^*$$

$$\sum_{j=1}^{n} \lambda_{4j} GW_j \geq \theta_{4k} GW_k$$

$$\sum_{j=1}^{n} \lambda_{4j} = 1$$

$$\lambda_{4j} \geq 0; \;\; j = 1,2, \ldots, n$$

$$\theta_{4k} \geq 0$$

Finally, we compute the number of games each MLB team would have won had it been efficient, i.e., the team's *efficient games won* $EGW_k = {}^*GW_k^* = \sum_{j=1}^{n} {}^*\lambda_{4j}^* GW_j$ where ${}^*\lambda_{4j}^*$ are the optimal weights obtained when solving the on-field integration model for team k, assuming the front office, the on-field offense, and the on-field defense are all efficient. The *organizational (overall team) inverse efficiency* for MLB team k is $\theta_k = {}^*GW_k^* / GW_k$.

3.1.3 NBA Network DEA Model

First, we solve the front office DEA model for NBA team k.

$$\text{Maximize } \theta_{1k}$$

$$\text{subject to}$$

$$\sum_{j=1}^{n} \lambda_{1j} SAL_j \leq SAL_k$$

$$\sum_{j=1}^{n} \lambda_{1j} TPG_j \geq \theta_{1k} TPG_k$$

$$\sum_{j=1}^{n} \lambda_{1j} TPS_j \leq \epsilon_{1k} TPS_k$$

$$\sum_{j=1}^{n} \lambda_{1j} = 1$$

$$\theta_{1k} + \epsilon_{1k} = 2$$

$$\lambda_{1j} \geq 0; \ \ j = 1,2, \dots, n$$

$$\theta_{1k}, \epsilon_{1k} \geq 0$$

Next, we solve the on-court offense and on-court defense DEA models for NBA team k.

$$\text{Maximize } \theta_{2k}$$

$$\text{subject to}$$

$$\sum_{j=1}^{n} \lambda_{2j} TPG_j \leq TPG_k$$

$$\sum_{j=1}^{n} \lambda_{2j} PG_j \geq \theta_{2k} PG_k$$

$$\lambda_{2j} \geq 0; \ \ j = 1,2, \dots, n$$

$$\theta_{2k} \geq 0$$

$$\text{Minimize } \varepsilon_{3k}$$

$$\text{subject to}$$

$$\sum_{j=1}^{n} \lambda_{3j} TPS_j \geq TPS_k$$

$$\sum_{j=1}^{n} \lambda_{3j} PS_j \leq \varepsilon_{3k} PS_k$$

$$\lambda_{3j} \geq 0; \;\; j = 1,2, \dots, n$$

$$\varepsilon_{3k} \geq 0$$

Finally, we solve the on-court integration DEA model for NBA team k.

$$\text{Maximize } \theta_{4k}$$

$$\text{subject to}$$

$$\sum_{j=1}^{n} \lambda_{4j} PG_j \leq PG_k$$

$$\sum_{j=1}^{n} \lambda_{4j} PS_j \geq PS_k$$

$$\sum_{j=1}^{n} \lambda_{4j} GW_j \geq \theta_{4k} GW_k$$

$$\sum_{j=1}^{n} \lambda_{4j} = 1$$

$$\lambda_{4j} \geq 0; \;\; j = 1,2, \dots, n$$

$$\theta_{4k} \geq 0$$

Let $TPG_k^* = \sum_{j=1}^{n} \lambda_{1j}^* TPG_j$ and $TPS_k^* = \sum_{j=1}^{n} \lambda_{1j}^* TPS_j$ where λ_{1j}^* is the optimal weights obtained when solving the front office model for team k. Now, we resolve the on-court offense model for team k using TPG_k^* as the RHS of the first constraint and resolve the on-court defense model for team k using TPS_k^* as the RHS of the first constraint.

$$\text{Maximize } \theta_{2k}$$

$$\text{subject to}$$

$$\sum_{j=1}^{n} \lambda_{2j} TPG_j \leq TPG_k^*$$

$$\sum_{j=1}^{n} \lambda_{2j} PG_j \geq \theta_{2k} PG_k$$

$$\lambda_{2j} \geq 0; \;\; j = 1,2, \dots, n$$

$$\theta_{2k} \geq 0$$

Minimize ε_{3k}

subject to

$$\sum_{j=1}^{n} \lambda_{3j} TPS_j \geq TPS_k^*$$

$$\sum_{j=1}^{n} \lambda_{3j} PS_j \leq \varepsilon_{3k} PS_k$$

$$\lambda_{3j} \geq 0; \;\; j = 1,2, \dots, n$$

$$\varepsilon_{3k} \geq 0$$

Let ${}^*PG_k^* = \sum_{j=1}^{n} {}^*\lambda_{2j}^* PG_j$ where ${}^*\lambda_{2j}^*$ are the optimal weights obtained when solving the on-court offense model for team k, assuming the front office is efficient and ${}^*PS_k^* = \sum_{j=1}^{n} {}^*\lambda_{3j}^* PS_j$ where ${}^*\lambda_{3j}^*$ are the optimal weights obtained when solving the on-court defense model for team k, assuming the front office is efficient. Finally, we resolve the on-court integration model for team k using ${}^*PG_k^*$ and ${}^*PS_k^*$ as the RHS of the first and second constraints, respectively.

Maximize θ_{4k}

subject to

$$\sum_{j=1}^{n} \lambda_{4j} PG_j \leq {}^*PG_k^*$$

$$\sum_{j=1}^{n} \lambda_{4j} PS_j \geq {}^*PS_k^*$$

$$\sum_{j=1}^{n} \lambda_{4j} GW_j \geq \theta_{4k} GW_k$$

$$\sum_{j=1}^{n} \lambda_{4j} = 1$$

$$\lambda_{4j} \geq 0; \;\; j = 1,2, \ldots, n$$

$$\theta_{4k} \geq 0$$

Finally, we compute the number of games each NBA team would have won had it been efficient, i.e., the team's *efficient games won* EGW_k ${}^*GW_k^* = \sum_{j=1}^{n} {}^*\lambda_{4j}^* \, GW_j$ where ${}^*\lambda_{4j}^*$ are the optimal weights obtained when solving the on-court integration model for team k, assuming the front office, the on-court offense, and the on-court defense are all efficient. The *organizational (overall team) inverse efficiency* for NBA team k is $\theta_k = {}^*GW_k^* / GW_k$.

3.2 Normal Approximation to the Binomial Distribution

We interpret the phrase *"reasonable hope of reaching post-season play"* to mean that a team must have at least the same probability of reaching the playoffs as it would have if all teams in its league were equally talented. We refer to this probability as the team's *balanced probability*. The balanced probability for a given team in a given season depends on the playoff qualification condition in effect. We interpret each probability as the minimum probability of qualifying for post-season play that a team must achieve to be competitive under its playoff qualification condition.

If all teams are equally talented, we can model the probability of winning a certain number of games in a season using a binomial distribution with the number of trials equal to the number of games in a season and a probability of success of 0.5 for each trial. Using the normal approximation to this binomial distribution and treating the balanced probabilities as upper-tail probabilities, we find MGW_k, the *minimum number of games that team k needs to win under each playoff qualification condition in order to be competitive*.

3.2.1 Playoff Qualification Conditions, Balanced Probabilities, and MGWs in MLB

From the 2001 to 2011 season, there were six teams in the NLC, five teams in the NLE, NLW, ALE, and ALC, and four teams in the ALW. Four teams (the three division winners and one wild card) made the playoffs from each league. For the 2012 season, the divisions remained the same as in 2001+ to 2011. However, two wild card teams made the playoffs from each league. From 2013 to 2018, each division consisted of five teams with the three division winners and two wild card teams making the playoffs

from each league. Thus, the balanced probability for a team in a given season depends on the number of teams in its league, the number of teams in its division, and the number of wild card teams in its league.

As an example, the balanced probability for an MLB team in a fifteen-team league with five teams in its division and two wild cards in its league is the sum of the probability the team wins the division and the probability the team wins a wild card given it does not win the division; $\frac{1}{5} + \frac{2}{12} \cdot \frac{4}{5} = \frac{12}{60} + \frac{8}{60} = \frac{20}{60} = \frac{1}{3}$. We model the probability of winning a certain number of games in a season using a binomial distribution with 162 trials and a probability of success of 0.5 for each trial. This distribution has a mean of 81 games and a standard deviation of 6.36 games. For a team in a fifteen-team league with five teams in its division and two wild cards in its league, $MGW_k = 83.7$. Similarly, we compute the balanced probability and MGW for each of the playoff qualification conditions.

Teams in League	Teams In Division	Wild Card(s	Balanced Probability	MGW	Season(s)	Division(s)
16	6	1	0.230769231	85.68589	2001-2011	NLC
16	5	1	0.261538462	85.06409	2001-2011	NLE NLW
14	5	1	0.272727273	84.84756	2001-2011	ALE ALC
14	4	1	0.318181818	84.00881	2001-2011	ALW
16	6	2	0.294871795	84.4315	2012	NLC
16	5	2	0.323076923	83.92177	2012	NLE NLW
14	5	2	0.345454545	83.53045	2012	ALE ALC
14	4	2	0.386363636	82.83797	2012	ALW
15	5	2	0.333333333	83.74113	2013-2018	All

3.2.2 Playoff Qualification Condition, Balanced Probability, and MGW in the NBA

Since the 2002-03 season, the NBA has 30 teams split into two 15 team conference. The eight teams with the highest winning percentage in each conference make the postseason. The balanced probability of any team making the playoffs is 0.5333. MGW is 40.62.

3.3 Gini Index

We then determine the minimum total player salary needed to be competitive in each season, which we call the competitive salary for that season. To do this, within each season, we sort the teams according to total player salary from low to high and use the Gini index to identify a total player salary that partitions the teams into two sets, one of which consists primarily of competitive teams and one of which consists primarily of noncompetitive teams. The competitive salary in that season is the total player salary of the lowest paid team in the primarily competitive set.

We now partition the noncompetitive teams into two groups:

- **Noncompetitive Due to Low Total Player Salary (NCS):** A noncompetitive team is *noncompetitive due to low total player salary* if its total player salary is less than the competitive salary.

An example of a team in this category is the 2002 Tampa Bay Devil Rays (see Figure 3.1.) The MGW for an AL East team in 2002 is 84.8. Since Tampa Bay's EGW is 55 in 2002, the team is noncompetitive. The 2002 MLB competitive salary is $38,670,500. Tampa Bay's total player salary in 2002 is $34,380,000. Thus, we conclude that Tampa Bay is noncompetitive due to low total player salary in 2002.

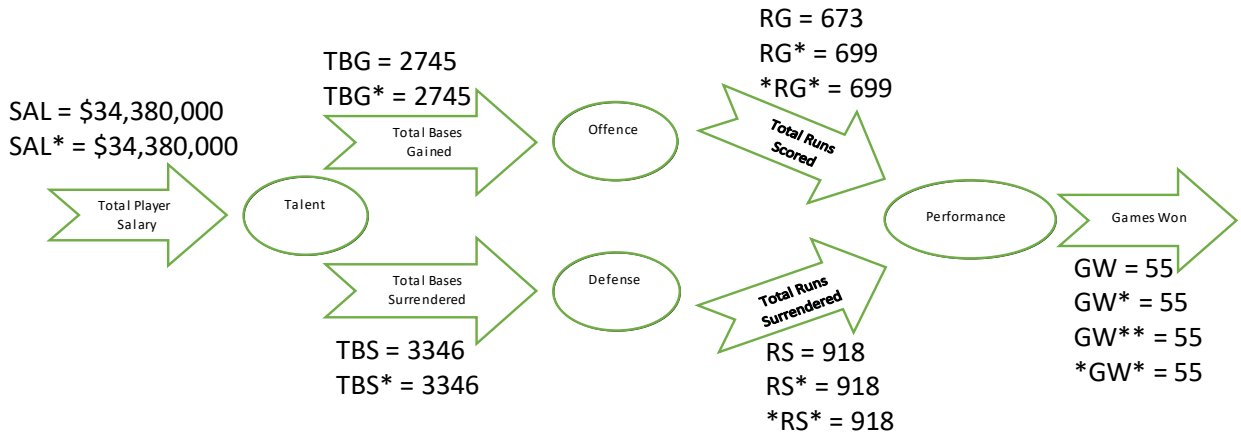


Figure 3.2: The 2002 Tampa Bay Devil Rays are noncompetitive due to low total player salary.

- **Noncompetitive for Other Reasons (NCO):** A noncompetitive team is *noncompetitive for other reasons* if its total player salary is greater than the competitive salary.

An example of a team in this category is the 2001 Colorado Rockies (see Figure 3.2.) The MGW for an NL West team in 2001 is 85.1. Since Colorado's EGW is 84 in 2001, the team is noncompetitive. The 2001 MLB competitive salary is $24,130,500. Colorado's total player salary in 2001 is $71,541,334. Thus, we conclude that Colorado is noncompetitive for other reasons in 2001.

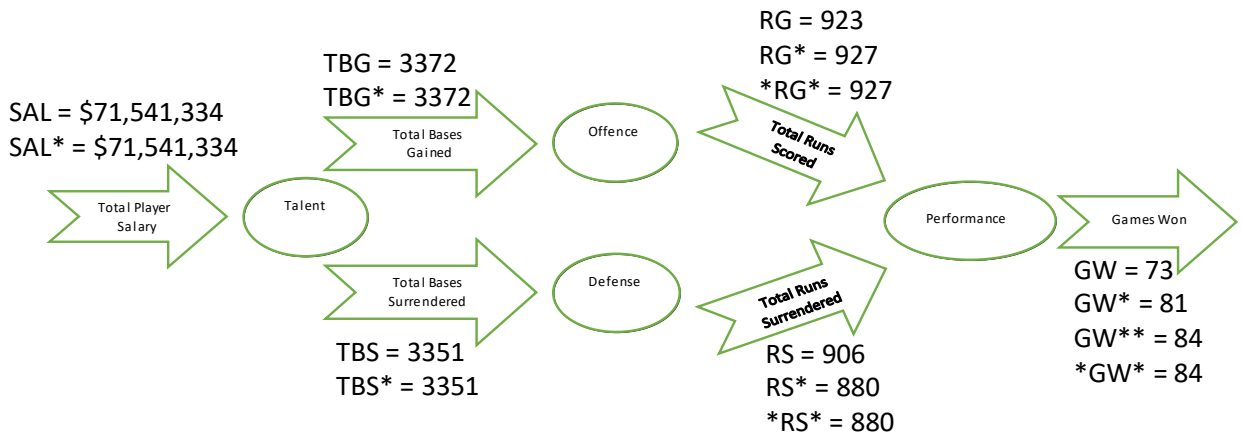


Figure 3.3: The 2001 Colorado Rockies are noncompetitive for other reasons.

Next, we analyze the competitive teams. To do this, we need to provide more definitions. Define GW^{**} to be the number of games that an efficient on-field operation would have won given the actual performance of the front office. We note that $GW \leq GW^{**} \leq EGW$. We obtain GW^{**} by performing a network DEA on the offense, defense, and integration DEA models. Let TPS^* be the total player salary of the efficient front office operation. We note that $TPS^* \leq TPS$. We obtain TPS^* from the front office DEA model.

We now partition the competitive teams into three groups:

- **Conditionally Competitive (CC)**: A competitive team is *conditionally competitive* if $GW^{**} < MGW$. The team is spending enough money on total player salary but inefficiency in the front office has resulted in insufficient player performance to win enough games to achieve the balanced probability of qualifying for post-season play. The front office must become more efficient for this to happen. We note that a conditionally competitive team may be overspending on player salaries if $TPS^* < TPS$.

An example of a team in this category is the 2001 Kansas City Royals (see Figure 3.3.) The MGW for an AL Central team in 2001 is 84.8. Since Kansas City's EGW is 112 in 2001, the team is competitive. However, the team's GW** is 80 in 2001. Thus, we conclude that Kansas City is conditionally competitive in 2001.

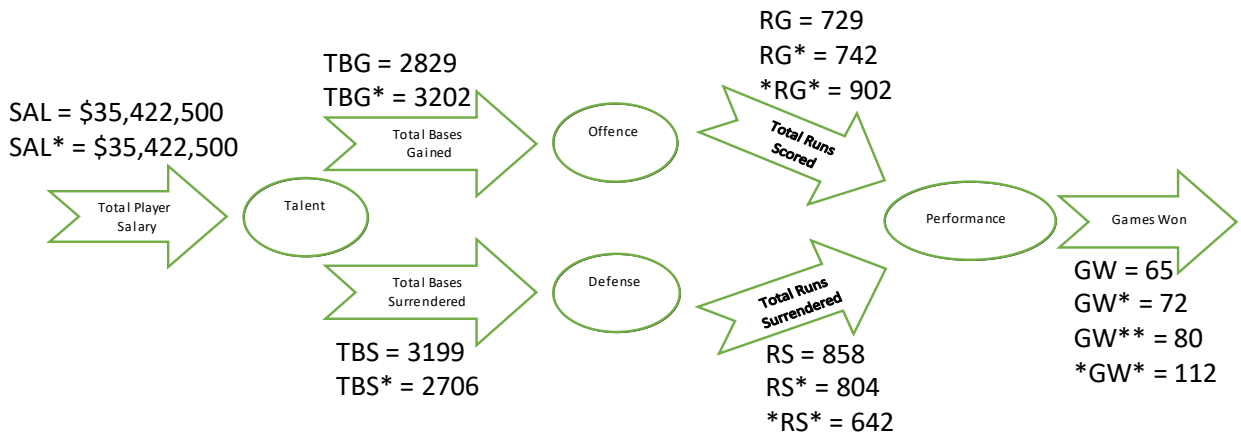


Figure 3.4: The 2001 Kansas City Royals are conditionally competitive.

- **Economically Competitive (EC)**: A competitive team is *economically competitive* if $GW^{**} \geq MGW$ and $TPS^{*} = TPS$. The team has sufficient player performance on the field to achieve the balanced probability of qualifying for post-season play. Moreover, there is no evidence that the team is overspending on total player salary.

An example of a team in this category is the 2001 Oakland Athletics (see Figure 3.4.) The MGW for an AL West team in 2001 is 84.0. Since Oakland's EGW is 116 in 2001, the team is competitive. Furthermore, the team's GW** is 100 in 2001 and there is no evidence that the team could reduce total player salary. Thus, we conclude that Oakland is economically competitive in 2001.

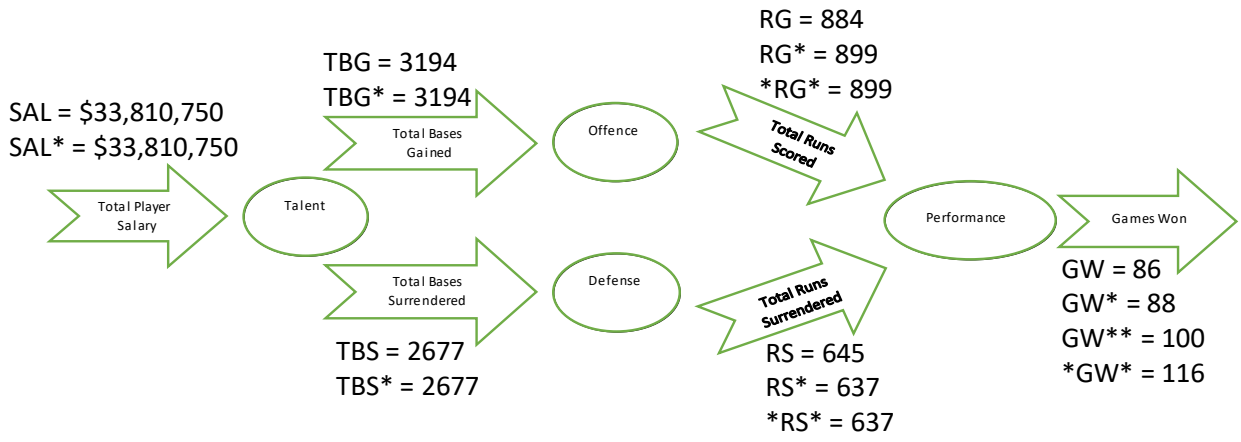


Figure 3.5: The 2001 Oakland Athletics are economically competitive.

- **Hypercompetitive (HC)**: A competitive team is *hypercompetitive* if $GW^{**} \geq MGW$ and $TPS^{*} < TPS$. The team has sufficient player performance on the field to achieve the balanced probability of qualifying for post-season play. However, there is evidence that the team is overspending on total player salary.

An example of a team in this category is the 2001 Los Angeles Dodgers (see Figure 3.5.) The MGW for an NL West team in 2001 is 85.1. Since Los Angeles' EGW is 112 in 2001, the team is competitive. Furthermore, the team's GW** is also 112 in 2001. However, there is evidence that the team could reduce total player salary from $109,105,953 to $74,720,834. Thus, we conclude that Los Angeles is hypercompetitive in 2001.

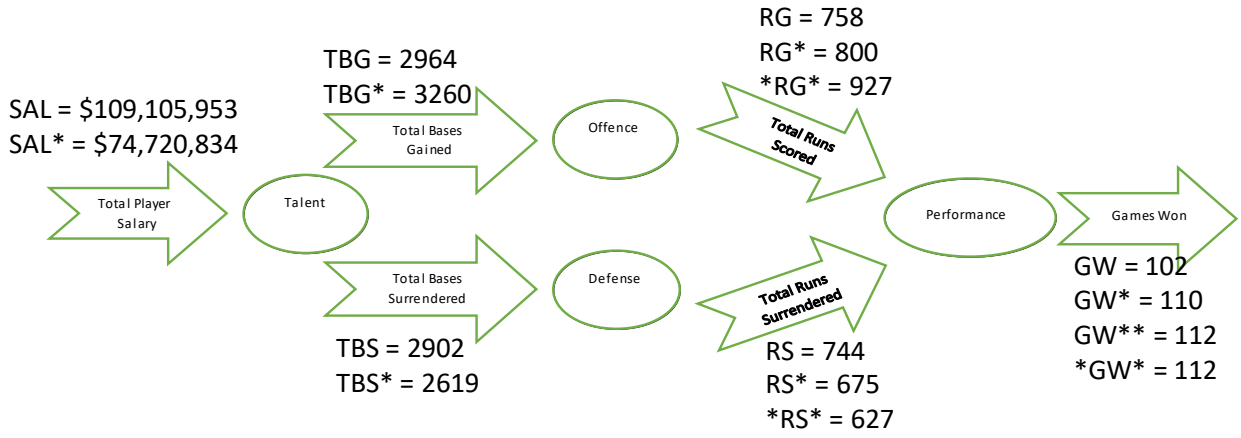


Figure 3.6: The 2001 Los Angeles Dodgers are hypercompetitive.

Finally, we use the Gini index again, this time to determine the value of total player salary that partitions hypercompetitive teams from other teams. We call this value of total player salary the *hypercompetitive salary* for the given season.

3.4 Summary of Methodology

Figure 3.6 presents an overview of the study methodology. In a given season, we apply network DEA to measure the front office, on-field (on-court), and overall team efficiency of each MLB (NBA) team. Next, we use the binomial distribution to classify teams as competitive versus noncompetitive. Finally, for each season, we use the Gini index to determine the minimum total player salary to be competitive and the maximum total player salary without overspending.

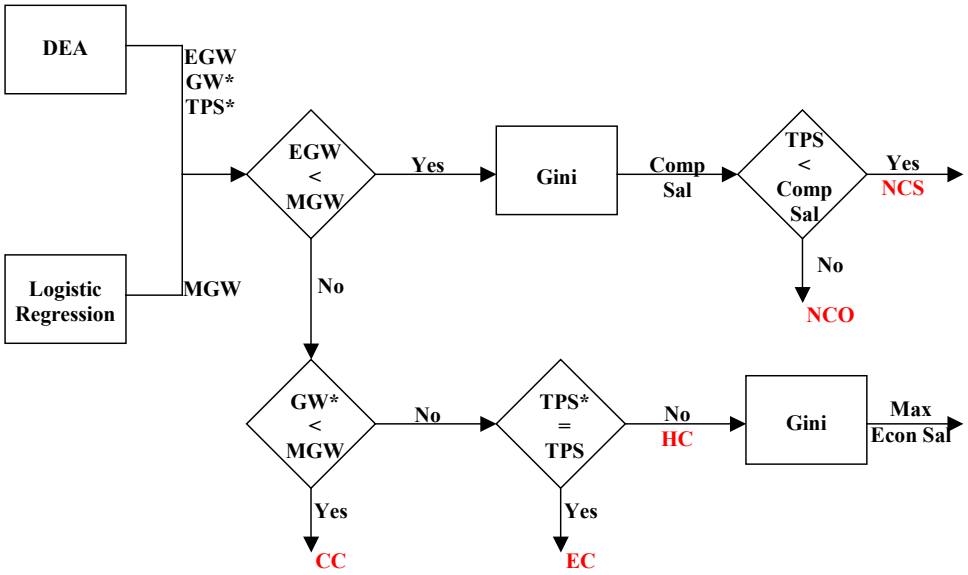


Figure 3.7: An overview of the methodology used in this study.

4. Data

The data for MLB and NBA team statistics are available from the Major League Baseball official website (MLB.com) and from the National Basketball Association official website (NBA.com), respectively. We were unable to find data on the number of opposition errors, which is required in the calculation of total bases gained. We estimated this value as a weighted average of the opposing teams' errors committed per game for each team in each season. The weights are the number of games the team played against each opponent. This approximation assumes that teams are equally likely to commit errors against each team they play.

The MLB and NBA official websites did not have a data for the total team salaries. We obtain this data from the USA Today website (USAToday.com) for MLB and from the Hoops Hype website (HoopsHype.com) for the NBA. We use the metropolitan area television market size for all teams in U.S. cities. For Toronto, we used the Neilson estimate.

5. Results

5.1 MLB Results

Twelve MLB teams (2.2%) are noncompetitive due to low total player salary. TB is NCS four times and MIA/FLA two times. Four MLB teams (0.7%) are noncompetitive due to other reasons. COL is NCO two times, LA one time, and TEX one time. 102 MLB teams (18.9%) are conditionally competitive. KC is CC nine times and BAL eight times. 319 MLB teams (59.1%) are economically competitive. OAK is EC in all

but one season. 103 MLB teams (19.1%) are hyp+ercompetitive. LA is HC twelve times and NYM, NYY, and LAA/ANA are each HC ten times. BOS, NYY, and STL are EC or HC in all seasons. Figure 5.1 shows MLB competitiveness by team and season. Figure 5.2 is a histogram of MLB competitiveness by team. Figure 5.3 is a histogram of MLB competitiveness by season.

Team Name	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
ARI	HC	HC	EC	CC	EC	EC	EC	EC	EC	CC	EC	EC	EC	CC	EC	CC	EC	EC
ATL	HC	HC	EC	HC	HC	EC	EC	EC	EC	EC	EC	HC	EC	EC	CC	EC	CC	EC
BAL	CC	CC	EC	EC	EC	CC	EC	EC	CC	CC	CC	EC	EC	EC	EC	EC	CC	CC
BOS	HC	HC	EC	EC	EC	EC	EC	EC	EC	EC	EC	HC	EC	HC	HC	EC	HC	EC
CHC	EC	EC	EC	HC	HC	CC	HC	EC	HC	HC	CC	CC	EC	EC	EC	EC	HC	HC
CHW	EC	EC	EC	EC	HC	EC	HC	HC	EC	HC	EC	HC	CC	EC	EC	EC	CC	CC
CIN	CC	EC	CC	CC	CC	EC	EC	EC	EC	EC	EC	HC	EC	EC	EC	CC	CC	CC
CLE	HC	EC	EC	EC	EC	EC	EC	EC	EC	CC	EC	CC	EC	EC	EC	EC	EC	EC
COL	NCO	CC	EC	CC	CC	EC	EC	EC	EC	EC	EC	CC	CC	NCO	CC	EC	EC	EC
DET	CC	CC	CC	EC	HC	EC	EC	HC	EC	EC	EC	HC	EC	EC	HC	HC	CC	CC
HOU	EC	EC	EC	EC	EC	HC	EC	EC	CC	CC	CC	CC	NCS	EC	EC	EC	EC	EC
KC	CC	CC	EC	CC	CC	CC	CC	CC	EC	CC	EC	EC	EC	EC	EC	EC	HC	CC
LA	HC	HC	NCO	HC	HC	EC	HC	HC	EC	HC	EC	HC	HC	HC	EC	HC	HC	EC
LAA/ANA	EC	EC	EC	HC	HC	HC	HC	HC	EC	HC	EC	HC	CC	EC	EC	EC	HC	HC
MIA/FLA	EC	EC	EC	EC	HC	EC	NCS	EC	EC	EC	EC	HC	NCS	EC	EC	EC	EC	CC
MIL	CC	CC	CC	NCS	EC	EC	EC	EC	CC	CC	EC	EC	EC	EC	EC	EC	EC	EC
MIN	EC	EC	EC	EC	EC	EC	HC	EC	EC	EC	CC	CC	CC	CC	EC	CC	EC	EC
NYM	HC	HC	CC	HC	HC	EC	EC	HC	HC	HC	EC	HC	EC	EC	EC	EC	CC	HC
NYY	HC	EC	EC	HC	EC	EC	EC	HC	EC	EC	HC	EC	HC	HC	HC	HC	HC	EC
OAK	EC	EC	EC	EC	EC	EC	EC	NCS	EC	EC	EC	EC	EC	EC	EC	EC	EC	EC
PHI	EC	EC	EC	EC	EC	EC	EC	EC	EC	HC	EC	HC	CC	HC	CC	NCS	EC	EC
PIT	CC	CC	EC	CC	EC	CC	EC	CC	CC	NCS	CC	EC	EC	EC	EC	EC	EC	EC
SD	EC	CC	EC	EC	HC	HC	EC	EC	EC	EC	EC	EC	EC	CC	EC	EC	NCS	EC
SEA	EC	EC	EC	EC	HC	EC	EC	CC	HC	CC	CC	CC	EC	EC	EC	EC	HC	EC
SF	EC	EC	EC	EC	CC	HC	HC	CC	EC	HC	EC	HC	HC	EC	EC	EC	CC	CC
STL	HC	EC	EC	EC	HC	EC	EC	EC	EC	HC	EC	HC	HC	EC	EC	EC	HC	EC
TB	CC	NCS	NCS	CC	NCS	CC	NCS	EC	EC	EC	EC	EC	EC	EC	EC	EC	EC	EC
TEX	CC	CC	CC	EC	EC	EC	EC	NCO	EC	EC	EC	EC	EC	CC	HC	EC	HC	CC
TOR	HC	EC	EC	EC	EC	EC	HC	EC	EC	EC	EC	CC	CC	EC	EC	EC	HC	CC
WAS/MON	CC	EC	EC	CC	CC	EC	EC	CC	CC	EC	EC	EC	HC	EC	EC	EC	HC	HC

Figure 5.1: MLB Competitiveness by Team and Season

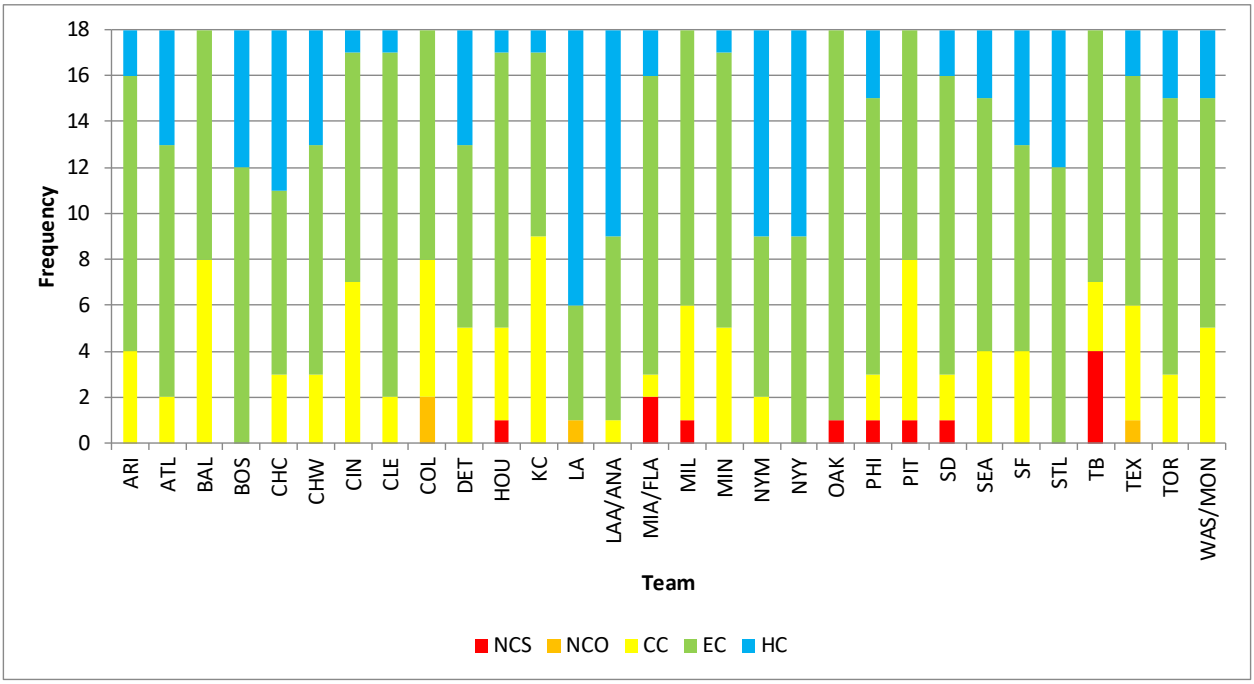


Figure 5.2: MLB Competitiveness Histogram by Team

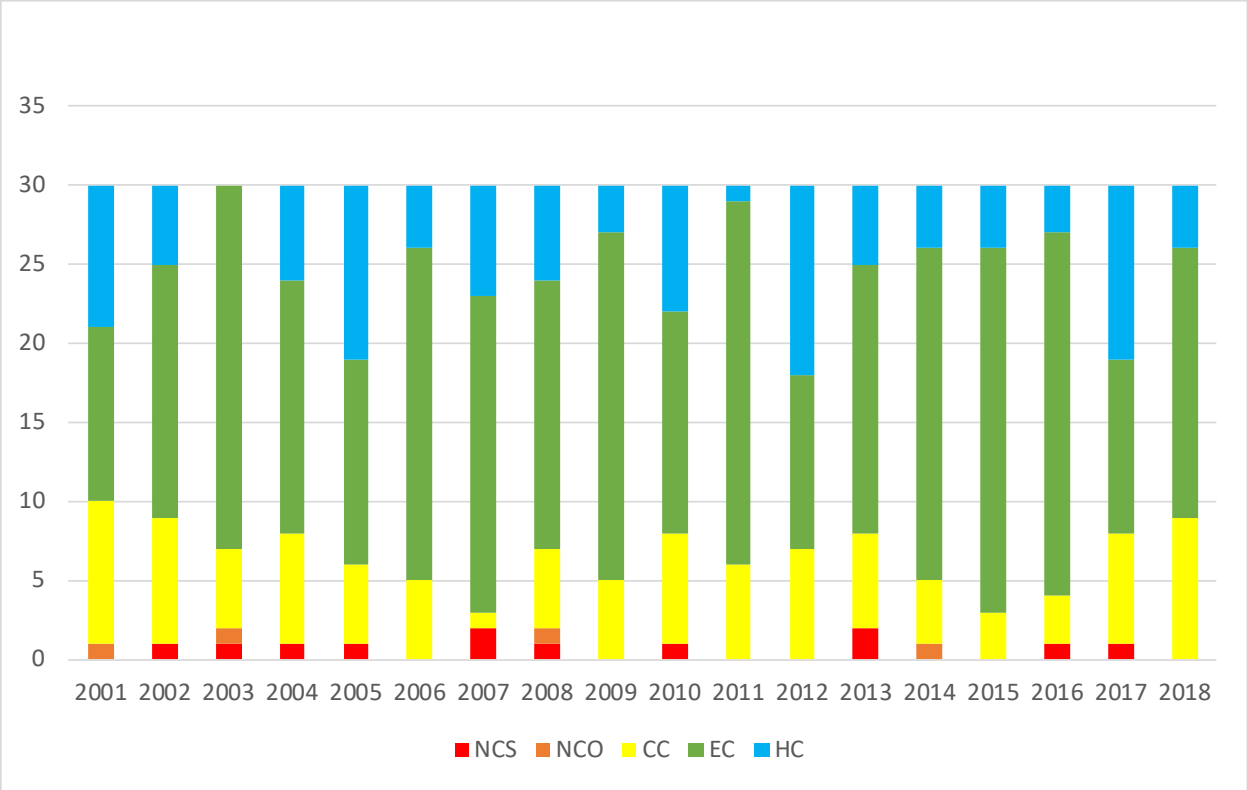


Figure 5.3: MLB Competitiveness Histogram by Season

5.2 NBA Results

Zero NBA teams are NCS, NCO, or CC. 246 NBA teams (58.6%) are EC The Pistons, Warriors, Clippers, Magic, and Spurs are each EC eleven times. 174 NBA teams (41.4%) are HC. The Knicks are HC eleven times, the Mavericks ten times, and the Lakers nine times. Figure 5.4 shows NBA competitiveness by team and season. Figure 5.5 is a histogram of NBA competitiveness by team. Figure 5.6 is a histogram of NBA competitiveness by season.

Team Name	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011	2011-2012	2012-2013	2013-2014	2014-2015	2015-2016	2016-2017	2017-2018
Atlanta Hawks	EC	EC	EC	EC	EC	HC	HC	HC	HC	EC	EC	EC	HC	EC
Boston Celtics	EC	EC	HC	EC	HC	HC	HC	HC	HC	EC	EC	EC	EC	EC
Charlotte Bobcats	EC	EC	EC	EC	EC	HC	EC	EC	HC	HC	HC	EC	EC	EC
Chicago Bulls	EC	EC	EC	EC	EC	HC	EC	EC	EC	HC	HC	EC	HC	EC
Cleveland Cavaliers	EC	EC	HC	HC	HC	HC	EC	EC	HC	HC	HC	EC	HC	EC
Dallas Mavericks	EC	EC	EC	HC	HC	HC	HC	HC	HC	EC	HC	HC	HC	HC
Denver Nuggets	EC	EC	HC	EC	EC	HC	HC	EC	HC	HC	HC	EC	HC	EC
Detroit Pistons	EC	EC	EC	EC	EC	HC	EC	HC	EC	HC	EC	EC	EC	EC
Golden State Warriors	EC	EC	HC	EC	EC	HC	HC	EC	EC	EC	EC	EC	EC	EC
Houston Rockets	EC	EC	HC	EC	EC	HC	HC	EC	HC	HC	EC	HC	HC	HC
Indiana Pacers	EC	HC	HC	EC	EC	HC	HC	EC	EC	EC	EC	EC	HC	HC
Los Angeles Clippers	EC	EC	EC	EC	EC	HC	EC	EC	HC	EC	HC	EC	EC	EC
Los Angeles Lakers	EC	EC	HC	EC	EC	HC	HC	HC	HC	HC	HC	HC	HC	EC
Memphis Grizzlies	EC	EC	HC	EC	EC	EC	HC	HC	HC	HC	HC	EC	HC	HC
Miami Heat	EC	EC	HC	HC	EC	HC	HC	HC	EC	EC	HC	EC	HC	EC
Milwaukee Bucks	EC	EC	HC	EC	EC	HC	HC	EC	HC	HC	HC	EC	HC	EC
Minnesota Timberwolves	EC	EC	HC	EC	EC	HC	EC	EC	EC	EC	EC	EC	HC	HC
New Jersey Nets	EC	EC	HC	EC	EC	HC	EC	EC	HC	EC	HC	EC	EC	EC
New Orleans Hornets	EC	EC	EC	EC	EC	HC	EC	EC	HC	EC	HC	EC	HC	HC
New York Knicks	EC	HC	HC	HC	HC	HC	HC	HC	HC	HC	HC	EC	HC	EC
Orlando Magic	HC	EC	EC	EC	EC	EC	HC	EC	HC	EC	EC	EC	EC	EC
Philadelphia 76ers	EC	HC	HC	EC	EC	HC	HC	HC	HC	HC	EC	EC	HC	EC
Phoenix Suns	EC	EC	HC	EC	EC	HC	HC	EC	HC	EC	EC	EC	EC	EC
Portland Trail Blazers	EC	EC	EC	HC	EC	HC	HC	HC	EC	EC	EC	EC	EC	EC
Sacramento Kings	EC	EC	HC	EC	EC	EC	HC	HC	HC	HC	HC	EC	HC	EC
San Antonio Spurs	EC	EC	EC	HC	EC	HC	EC	EC	EC	EC	HC	EC	EC	EC
Seattle SuperSonics	EC	EC	EC	EC	EC	HC	HC	HC	HC	EC	HC	EC	HC	EC
Toronto Raptors	EC	EC	EC	HC	EC	HC	HC	EC	HC	HC	HC	EC	EC	EC
Utah Jazz	EC	EC	EC	EC	EC	HC	HC	EC	HC	EC	EC	EC	HC	EC
Washington Wizards	EC	EC	HC	EC	EC	HC	HC	EC	EC	HC	HC	EC	HC	HC

Figure 5.4: NBA Competitiveness by Team and Season

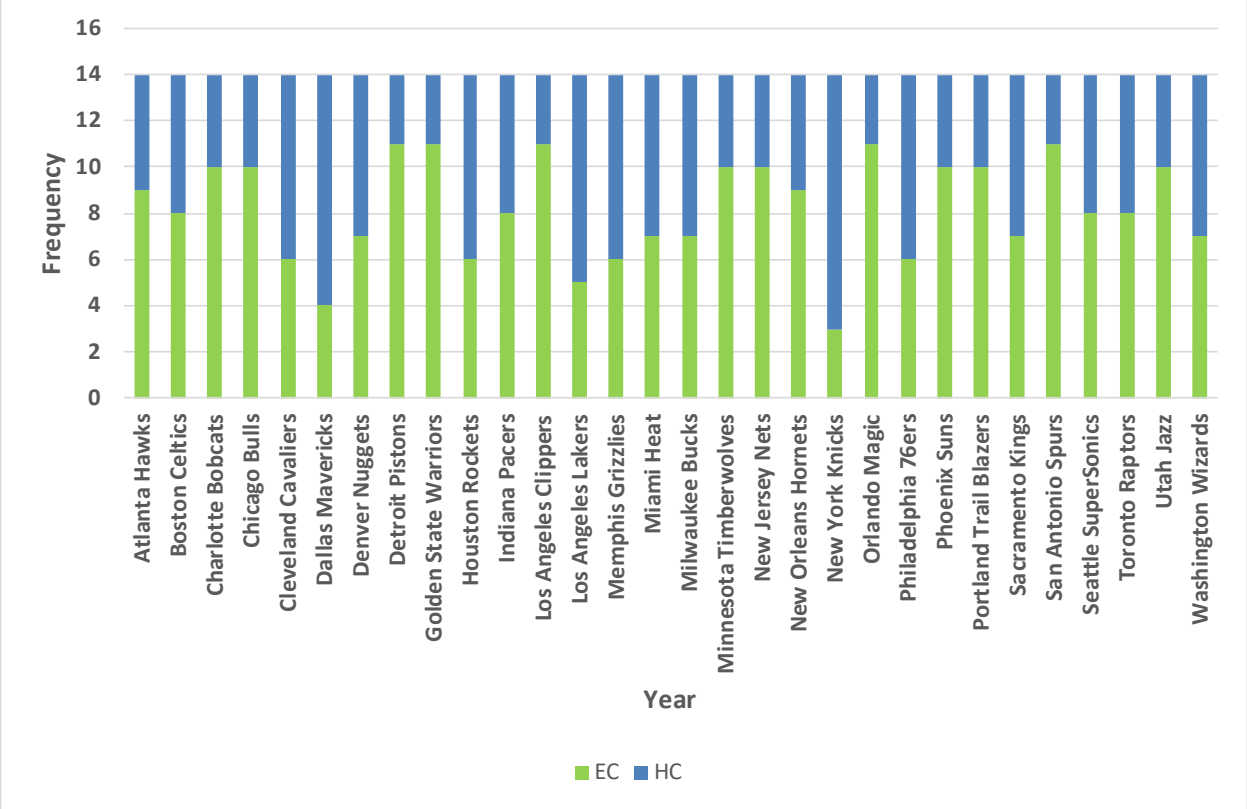


Figure 5.5: NBA Competitiveness Histogram by Team

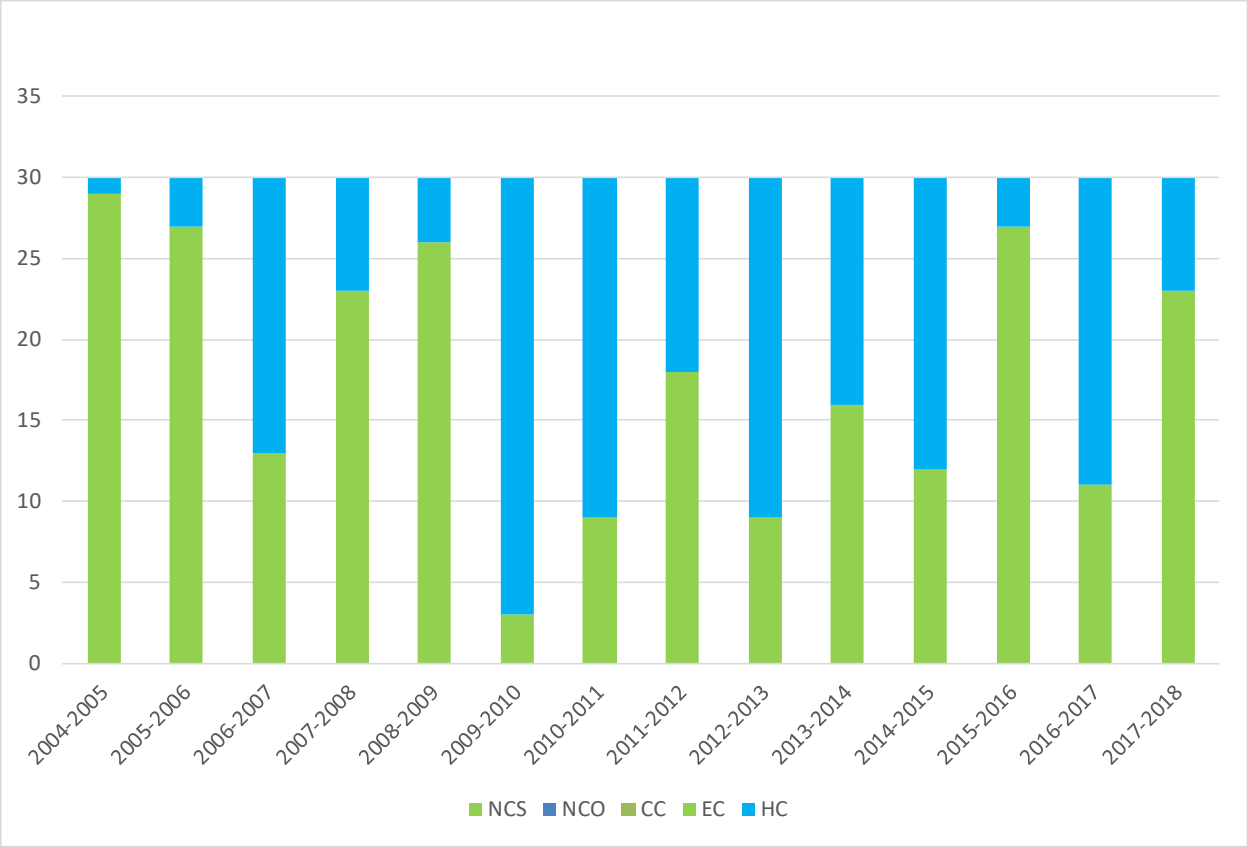


Figure 5.6: NBA Competitiveness Histogram by Season

6. Research Question Discussion

1. How many MLB and NBA teams are noncompetitive due to low total player salary?

Throughout the study period, there are twelve MLB teams that are noncompetitive due to salary and there are zero NBA teams that are noncompetitive due to low total player salary. However, it does not appear to be a systematic problem for any of the teams and no team is consistently NCS.

Season	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
NCS	0	1	1	1	1	0	2	1	0	1

2011	2012	2013	2014	2015	2016	2017	2018	Total NCS Teams
0	0	2	0	0	1	1	0	12

2. How many MLB and NBA teams are overspending on total player salary?

The number of MLB and NBA teams overspending on total player salary varies greatly from year to year. But on average, as we predicted, most of the teams overspending on team salary are from the larger markets such as New York, Texas, and Los Angeles.

MLB

Season	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
HC	9	5	0	6	11	4	7	6	3	8

2011	2012	2013	2014	2015	2016	2017	2018	Total HC Teams
1	12	5	4	4	3	11	4	103

NBA

Season	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011
HC	1	3	17	7	4	27	21

2011-2012	2012-2013	2013-2014	2014-2015	2015-2016	2016-2017	2017-2018	Total HC Teams
12	21	14	18	3	19	7	174

3. How much does an MLB and an NBA team need to spend on total player salary to be competitive?

On average the competitive salary in MLB has been increasing since 2001 but it varies from season to season. There have been seasons where it has decreased from or stayed the same as the previous season. In general, the competitive salary hovers around the minimum team salary, indicating that a team need not spend a large amount of money to compete in MLB.

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009
Competitive	$24,130,000	$38,670,500	$40,518,000	$29,556,667	$36,881,000	$14,998,500	$37,347,500	$48,689,783	$36,834,000

2010	2011	2012	2013	2014	2015	2016	2017	2018
$38,507,633	$34,366,000	$49,372,500	$39,199,500	$44,985,800	$63,564,200	$56,616,076	$49,143,998	$62,652,500

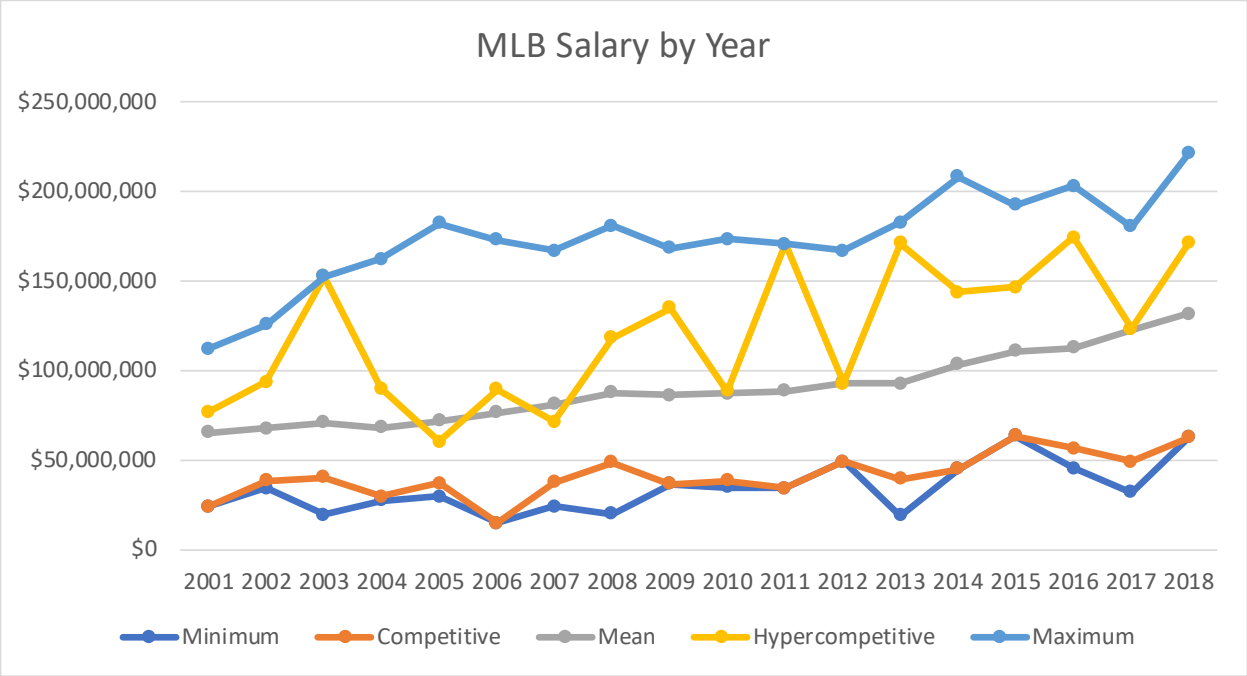


Figure 6.1: MLB Team Salary by Season

For the NBA, since all teams were competitive, the minimum competitive salary is just equal to the minimum salary that any team spent in a season. The trend is similar to that in MLB.

Year	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11
Competitive	$ 23,922,578.00	$ 33,458,932.00	$ 41,964,046.00	$ 53,745,468.00	$ 55,958,148.00	$ 56,946,258.00	$ 45,268,465.00

2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18
$ 48,950,236.00	$ 54,317,436.00	$ 53,308,029.00	$ 56,219,939.00	$ 62,488,858.00	$ 68,394,193.00	$ 85,147,034.00

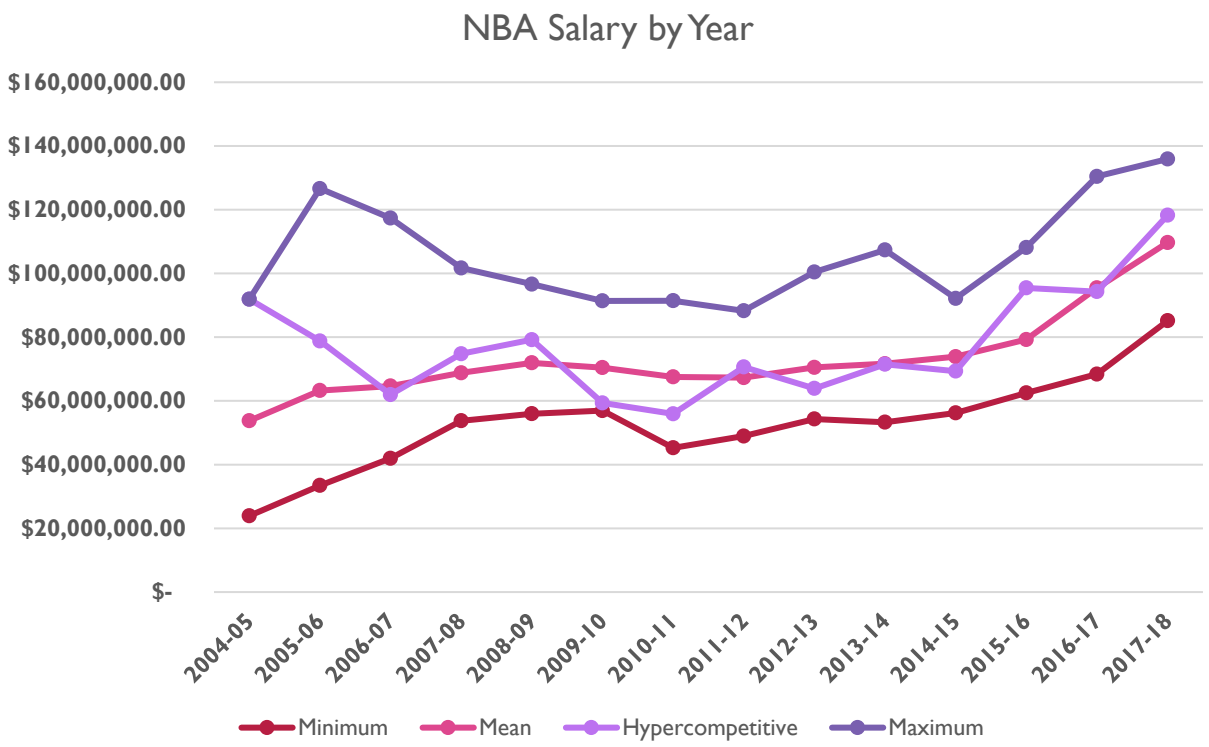


Figure 6.2: NBA Team Salary by Season

4. What is the maximum total player salary that an MLB and an NBA team can pay without overspending?

On average the hypercompetitive salary has been increasing over time but it varies from season to season. There have been seasons where it has decreased from or stayed the same as the previous season. In MLB, the hypercompetitive salary is typically above the mean team salary, while in the NBA, it hovers around the mean team salary.

MLB

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Hypercompetitive	$76,895,999	$93,470,367	$152,749,814	$90,182,500	$60,408,834	$89,728,419	$71,439,500	$118,164,036	$134,809,000	$88,817,251

2011	2012	2013	2014	2015	2016	2017	2018
$170,689,028	$92,834,433	$171,050,552	$143,862,395	$146,494,748	$174,330,522	$123,549,166	$171,248,332

NBA

Year	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10
Hypercompetitive	$ 91,908,443	$ 78,791,779	$ 61,960,489	$ 74,767,603	$ 79,188,973	$ 59,366,715

2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18
$ 55,922,488	$ 70,628,007	$ 63,901,224	$ 71,488,484	$ 69,322,836	$ 95,454,038	$ 94,287,153	$ 118,271,488

5. How does noncompetitiveness due to low total player salary relate to market size in MLB and in the NBA?

There appears to be a slight inverse relationship between market size and noncompetitiveness due to salary reasons in MLB. This means that the larger the market size of the team, the less likely they are to be NCS. This makes sense since large market teams generally have more revenue to spend and are a more attractive destinations for free agents. We note that all of the MLB teams that are NCS are in markets with less than 3.251 million TV homes. TB is most frequently NCS (four times) and is in the third smallest TV market. We also see that none of the big market teams (NYM, NYY, LA, LAA/ANA, etc.) are NCS even once during this study period.

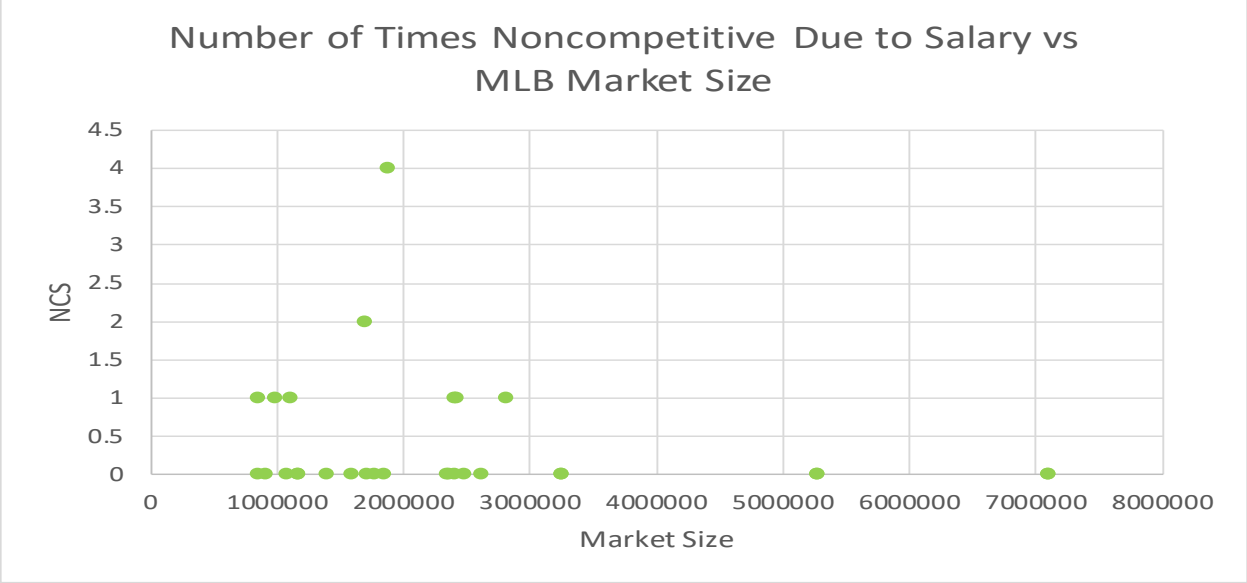


Figure 6.3: MLB Noncompetitiveness vs. Market Size

6. **How does overspending on total player salary relate to market size in MLB and in the NBA?**

In MLB, there is a positive association between market size and overspending on player salary. We see that LA is HC twelve times and NYM, NYY, and LAA/ANA are each HC ten times during the study period.

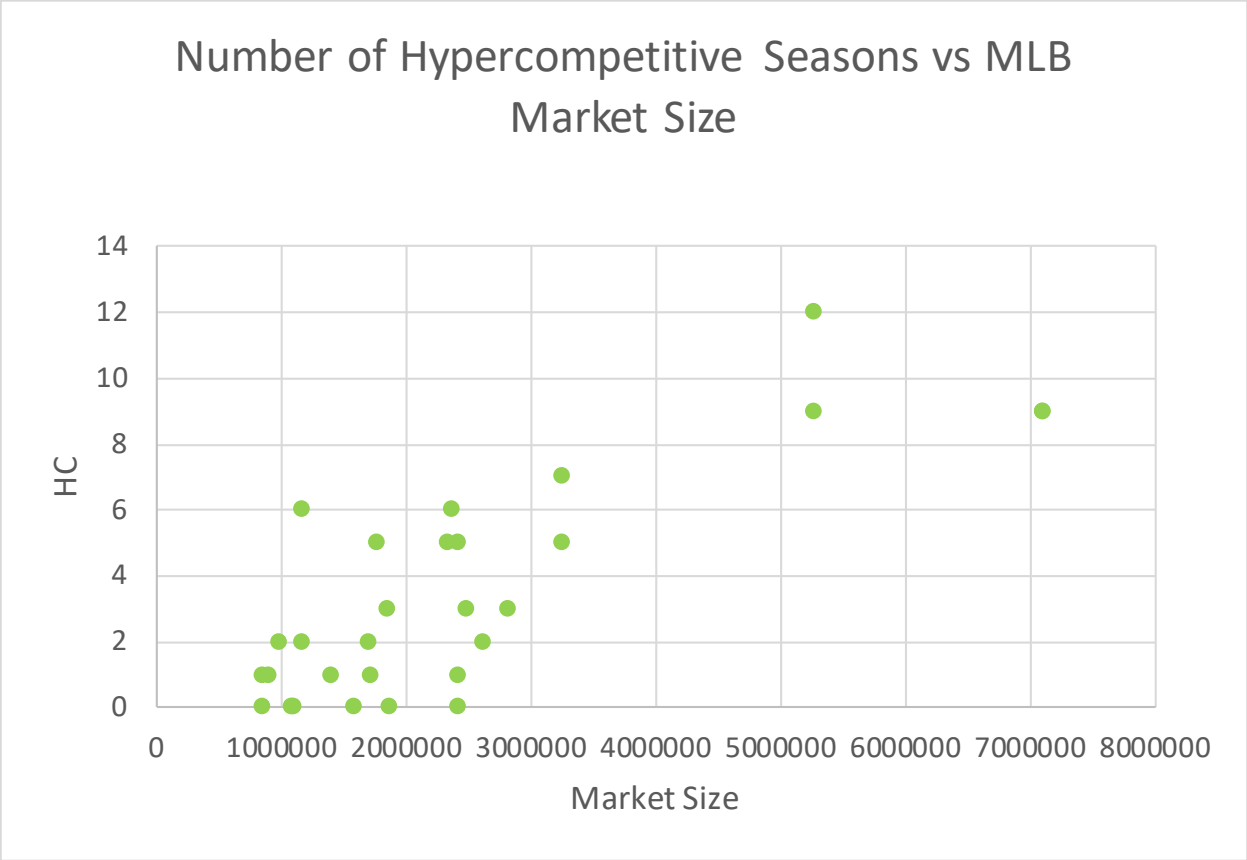


Figure 6.4: MLB Hypercompetitiveness vs. Market Size

The relationship is not as prevalent in the NBA, but a slight trend does exist. We notice that the Knicks and Lakers are overspending with a higher frequency than the Nets and Clippers even though they are in the same market. Therefore, we hypothesized that it is because being in a the same market as a much more storied and historic franchise takes away attention and potential free agents from the younger team.

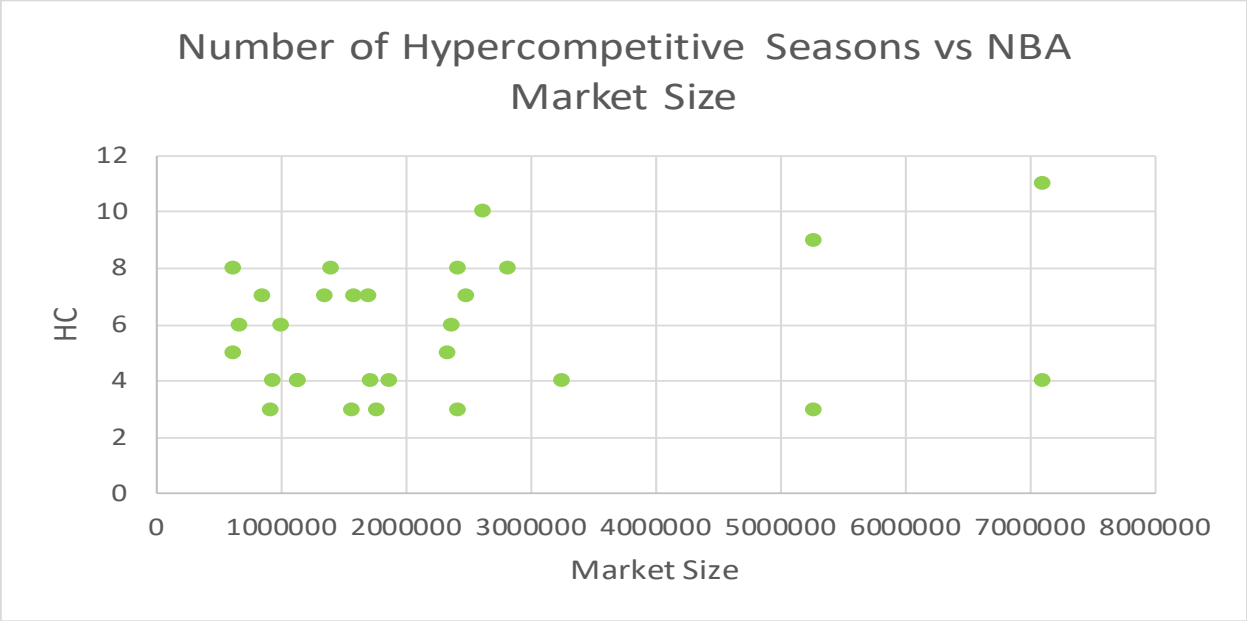


Figure 6.5: NBA Hypercompetitiveness vs. Market Size

7. Is there competitive balance in MLB and in the NBA?

MLB has sixteen teams that were noncompetitive due to salary or other reasons. In addition, there are 102 conditionally competitive teams in MLB. On the other hand, there appears to be competitive balance in the NBA. For every season there are zero noncompetitive or conditionally competitive teams. Thus, no team is unable to compete due to the amount of salary spent. We attribute this competitive balance to the existence of a harsh luxury tax in the NBA that punishes teams for going above the salary cap limit. There is also a hard cap, a dollar limit that no team is allowed to go over, even after paying a luxury tax. No such hard cap exists in MLB.

7. Conclusion

Overall, we believe there exists competitive balance within the two leagues. Noncompetitiveness due to salary reasons does not seem to be a significant problem in MLB or the NBA. In MLB, 422 (78.2%) of the teams in the analysis are classified as either economically competitive or hypercompetitive. In the NBA, every team is classified as either CC or HC. Although there exist certain instances of MLB teams being noncompetitive due to salary reasons, it does not appear to be a systematic problem. Furthermore, no MLB team is consistently noncompetitive. TB is NCS the most (four of the 18 seasons) and has not been NCS since 2007. It is safe to say that there exists a positive correlation between market size and competitiveness in MLB and the NBA. This correlation is smaller in the NBA.

In terms of market size and total team salary, there exists a negative relationship between the number of times an MLB team is noncompetitive due to low salary and the market size in which the team plays. This means that larger market teams are less likely to be noncompetitive due to salary reasons than their smaller market counterparts. There also exists a positive relationship between the number of times an MLB team is hypercompetitive and the market size in which the team plays. This indicates that MLB teams in larger markets are more likely to overspend on player salary. Although we see that large market NBA teams such as the Lakers and the Knicks are frequently hypercompetitive, there is not enough evidence to state that there is a relationship between the market size of the team and the number of times that it is hypercompetitive.

We attribute any disparity in competitive balance between MLB and the NBA to two factors. First, although there exists a soft cap in both leagues, there is still a hard cap in the NBA that is not present in MLB. This means that in both leagues, teams have to pay a fine for exceeding the salary cap. However, in the NBA, there is a set number that a team may not exceed. There is no such limit that exists in MLB. This is evident in the fact that the average maximum – minimum salary in MLB is around $139 million while it is only around $53 million in the NBA. The second factor is the number of teams that make the post season in the two leagues. Sixteen of the thirty teams make the post season in the NBA, 8 from each conference. In MLB only twelve of the thirty teams, six from each league qualify.

Average	**MLB**	**NBA**
Maximum – Minimum	$139,310,633	$ 52,834,590.21
Competitive - Minimum	$7,145,546	N/A
Maximum - Hypercompetitive	$53,015,601	$ 28,179,583.07

References

Anderson, T.R. & Sharp, G.P. (1997). A new measure of baseball batters using DEA. Annals of Operations Research 73: 141.

Bolen, J. B., Rezek, J., & Pitts, J. D. (2019). Performance Efficiency in NCAA Basketball. Journal of Sports Economics, 20(2), 218–241.

Carlos Pestana Barros & Stephanie Leach (2006). Performance evaluation of the English Premier Football League with data envelopment analysis. Applied Economics, 38:12

Cooper, W.W. & Ruiz, José L. & Sirvent, Inmaculada (2009). Selecting non-zero weights to evaluate effectiveness of basketball players with DEA. European Journal of Operational Research. 195. 563-574.

Einolf, K. W. (2004). Is Winning Everything?: A Data Envelopment Analysis of Major League Baseball and the National Football League. Journal of Sports Economics, 5(2), 127–151

Gutiérrez, Ó., & Ruíz, J.A. (2013). Data Envelopment Analysis and Cross-Efficiency Evaluation in the Management of Sports Teams: The Assessment of Game Performance of Players in the Spanish Handball League

HoopsHype.com https://hoopshype.com/salaries/

Lewis, H. F., Lock, K. A., & Sexton, T. R. (2007). Player Salaries, Organizational Efficiency, and Competitiveness in Major League Baseball. Journal of Sports Economics 266

Lewis, H. F., Lock, K. A., & Sexton, T. R. (2009). Organizational capability, efficiency, and effectiveness in Major League Baseball: 1901–2002. European Journal of Operational Research

Lewis, H. F., & Sexton, T. R. (2004). Data envelopment analysis with reverse inputs and outputs. Journal of Productivity Analysis, 21(2), 113–132.

Lozano, S., Villa, G., Guerrero, F., & Cortés, P. (2002). Measuring the performance of nations at the Summer Olympics using data envelopment analysis. Journal of the Operational Research Society, 53:5, 501-511.

MLB.com http://mlb.mlb.com/stats/sortable

NBA.com https://stats.nba.com/teams/traditional/?sort=W_PCT&dir=-1

Porter, P.K., & Scully, G.W. (1982). Measuring Managerial Efficiency: The Case of Baseball. Southern Economic Journal, Vol. 48, No. 3,pp. 642-650

Plácido Moreno, Sebastián Lozano (2014). A network DEA assessment of team efficiency in the NBA: 2014. Annals of Operations Research, Volume 214, Number 1, Page 99

Radovanovic, Sandro & Radojicic, Milan & Savic, Gordana (2014). Two-phased DEA-MLA approach for predicting efficiency of NBA players. Yugoslav Journal of Operations Research.

Ruggiero, J. (2011). Frontiers in Major League Baseball: Nonparametric Analysis of Performance Using Data Envelopment Analysis. New York, NY: Springer Science Business Media, LLC.

USAToday.com https://www.usatoday.com/sports/mlb/salaries/2018/team/all/

EFFICIENCY-BASED FLEET LOCATION-ALLOCATION-ROUTING SCHEME DESIGN UNDER THE RISK OF ROUTE DISRUPTIONS

TRACK: OPERATIONS MANAGEMENT/OPERATIONS RESEARCH

Competitive Paper Submitted for Presentation at the 52nd Annual Meeting of the NEDSI, Washington, DC, March 30-April 1, 2023

EFFICIENCY-BASED FLEET LOCATION-ALLOCATION-ROUTING SCHEME DESIGN UNDER THE RISK OF ROUTE DISRUPTIONS

ABSTRACT

The paper applies the goal programming (GP) approach to the fleet location-allocation-routing scheme design problem under the risk of route disruptions, given a set of transportation requests. All routes are susceptible to disruptions. For successful mission completion, three performance measures are considered: total routing distance, total routing distances, and the probability of routing all given sites without being disrupted. Solving the GP model with various sets of weights assigned to each measure will generate many different schemes. The data envelopment analysis methods are applied to evaluate those generated schemes. A case study demonstrates the applicability of the proposed procedure.

Keywords: Goal programming, Transshipment scheme, Route disruption, Data envelopment analysis,

INTRODUCTION

This paper investigates the fleet location-allocation-routing (FLAR) scheme design problem under the risk of route disruptions. A vehicle should visit some designated sites using routes from the initial location site and returning to the site. Each route is subject to disruptions with a certain risk probability. This study considers the efficiency, risk, and robustness of responding to a given routing mission. Two kinds of decision plans are usually considered for the FLAR scheme decisions. One is a strategic decision plan on the location. At the same time, the other is an operational decision plan on allocating the fleet to the sites and the routing decision. Daskin (2013) insists that the routing decision includes (i) which sites to assign to which routes and (ii) in what order sites should be served on each route. The traditional models for strategic design focused primarily on minimizing the related costs, ignoring the fact that the fleets' routes are at risk of disruptions.

As several references cited in Farahani et al. (2010, 2015), Fang and Li (2015), and Hong and Jeong (2019) demonstrate, location-allocation problems are fundamentally multi-objective, where those objectives frequently conflict with each other. The multi-objective programming

(MOP) technique provides an analytical framework where various objectives can be focused on simultaneously so that a decision-maker can provide optimal solutions. For efficient FLAR scheme design, this study formulates the network design problem with three objectives under the risk of route disruption as a goal programming (GP) model, which is a widely used MOP model. Contrary to a single objective optimization problem that can define the best solution, the notion of the best solution does not exist in the MOP model. Most MOP techniques require the decision-maker's judgment to provide weights assigned to the deviational variables from the target values in the objective function. The decision-maker appropriately reflects the desirability of deviations based on the importance of the various target values. As the number of goals or objectives increases, solving the GP model will yield many alternative options because each weight factor set for the goals generates a different option. Similarly, it would be possible for the decision-makers to develop many FLAR schemes, changing the values of those weights.

Evaluating various alternatives and identifying the most efficient options would be essential to efficient schemes. It would be difficult to change the fixed network scheme later, which is related to FLAR decisions. There is no standard procedure available to assign weight values in a way that the decision-makers find the most desirable solution. Evaluating alternatives generated by solving the GP model can be considered a multiple-criteria decision-making (MCDM) problem, requiring a systematic evaluation system. The question is how to evaluate these network schemes without any biased preferences or decision makers' subjective judgment and how to identify the most efficient FLAR network schemes for the decision-makers to consider adopting.

Data envelopment analysis (DEA) is one of the widely used methodologies to assess the efficiency of decision-making units (DMUs) with multiple inputs to use and outputs to produce. The classical DEA (C-DEA), developed by Charnes et al. (1978), generates a single, comprehensive performance measure for each DMU. The best ratio would identify the most efficient DMU among all the DMUs. C-DEA, with its self-evaluation, forces each DMU to be evaluated with its most favorable weights, ignoring unfavorable inputs/outputs to maximize self-efficiency. Consequently, it has suffered mainly from a lack of discrimination since it generates too many efficient DMUs. The two most popular methods for remedying C-DEA deficiency are the cross-efficiency DEA (CE-DEA) and the supers-efficiency DEA (SE-DEA) methods. This

paper applies C-, CE, and SE-DEA methods to evaluate and rank the FLAR network schemes generated by solving the GP model

This paper is organized as follows. After the literature review of the multi-objective FLAR models in the following section, we provide a brief background for the FLAR model with GP. Then, the proposed method of combining a GP model and DEA techniques is discussed. Next, we demonstrate the proposed method by GP model formulation and DEA evaluation through a case study, followed by conclusions.

LITERATURE REVIEW

This study is motivated by Klimberg and Ratick (2008) and Fang and Li (2015), who consider the facility location-allocation problem and apply the DEA method to find the optimal location-allocation in terms of efficiency. Klimberg and Ratick (2008) assume that locating facilities at different potential sites may affect the performance of the facility's capability to convert inputs into outputs. Based on this postulation, they develop and test facility location modeling formulations by applying DEA. Following Kimberg and Ratick (2008), Fang and Li (2015) present a multiple-objective linear programming model where the DEA method is integrated with the facility location-allocation problem. But these two approaches assume that both inputs and outputs are pre-fixed for DEA to be used. The model of Klimberg and Ratick (2008) requires a considerable amount of pre-determined input and output data. Consequently, as the number of facilities and potential sites increase, many constraints for their simultaneous DEA model are required. Hong and Jeong (2019) consider a similar problem with a different approach. Contrary to the pre-fixed inputs and outputs, they (2019) generate inputs and outputs by solving a multi-objective programming model, then apply the DEA methods for the generated inputs and outputs.

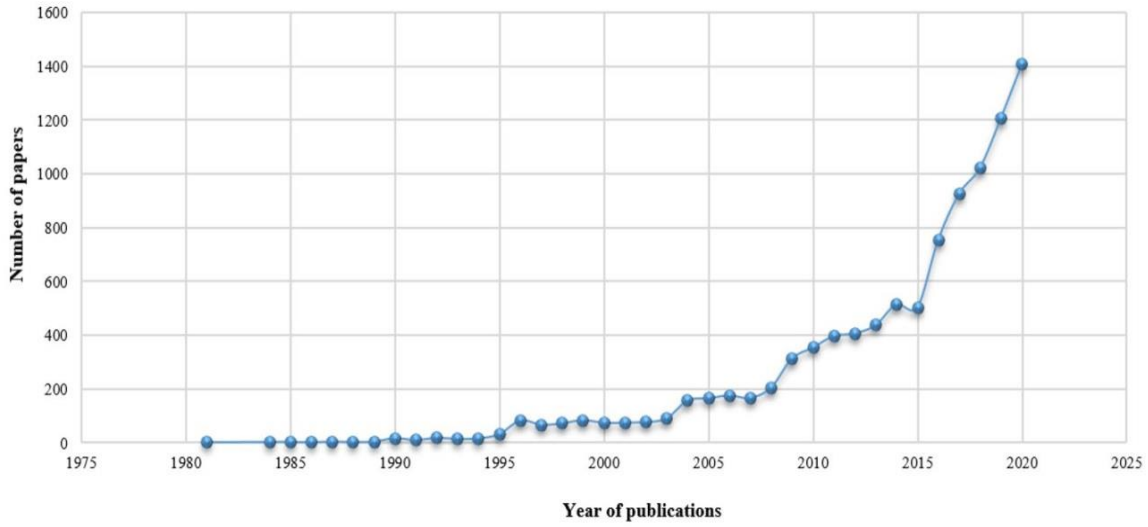
Data envelopment analysis (DEA) is a widely used methodology to rate a set of peer organizations. In the context of DEA, these organizations are called decision-making units (DMUs), which use inputs to produce outputs. The DEA method has been widely accepted as an effective technique since the DEA models need not follow the exact behavior function of those DMUs regarding transforming inputs to outputs. The non-parametric approach solves a linear program (LP) formulation for each DMU, and the weights assigned to each linear aggregation result from the corresponding LP. The DMUs under evaluation should be relatively

homogeneous for the DEA to measure relative efficiency, where the general expression for efficiency is the ratio of outputs to inputs. If there are multiple outputs or inputs, the problems of using C-DEA would be apparent. As the number of inputs and outputs of DMUs increases, these issues become more critical in evaluating and ranking them. Meza & Jeong (2013) and Ramanathan (2006) insist that a considerably large set of DMUs is required for the assessment to be meaningful, as the entire technique is grounded on a comparison of each DMU with all the other ones. DEA eventually determines which DMUs make efficient use of their inputs and produce the most outputs and which do not. Thus, the DEA models classify DMUs into two groups, efficient and inefficient DMUs. The DEA method can quantify what improved performance levels should be attainable for inefficient DMUs. Also, the analysis indicates where an inefficient DMU might search for benchmarking help as it looks for ways to improve.

Due to its enhanced discriminating power, many applications based on the CE evaluation have been published (see Sexton et al., 1986; Paryzad et al., 2018; Lee, 2019; Liu et al., 2019). As Doyle and Green (1994) note, the primal issue is the non-uniqueness of CE scores due to the often-present multiple optimal DEA weights. The second issue is that the CE method frequently ranks inefficient DMUs ahead of fully efficient DMUs. The idea of super-efficiency (SE), developed by Anderson and Peterson (1993), is that the C-DEA model is applied, excluding a DMU under evaluation from the reference set of the C-DEA model. Charnes et al. (1992) use the SE-DEA model to study the sensitivity of the efficiency classification. Anderson and Peterson (1993) propose the SE model for ranking efficient DMUs. But the critical issue of using the model is that the adjacent DMUs decide the SE score (SES) of an efficient DMU, so it would sometimes be unreasonable for some DMUs, significantly lowered ranked DMUs, to be ranked by the SESs.

As Zhu (2022) points out, the various DEA-based methods have been developed and widely applied in many different areas to evaluate the relative performance of DMUs, resulting in **more than 5,000 publications in the Web of Science database**. Panwar et al. (2022) also observe that from 1978 to 1995, there has been a constant growth in the publications of DEA. Still, from 1995 onwards, there has been an exponential rise in theoretical development and diverse applications. See Figure 1, which demonstrates such a trend for DEA publications since 1980.

Figure 1. DEA publications since 1980



GOAL PROGRAMMING FORMULATION

The following nomenclature is used:

Sets:

M: index set of potential fleet sites *(j, k, i =1, 2, …, M)*

Parameters:

b_j: minimum number of sites that fleet *j* should cover

B_j: maximum number of sites that fleet *j* can cover

d_{ik}: routing distance between site *i* and site *m*

t_{ik}: routing time between site *i* and site *m*

p_{ik}: the risk probability of the route's being disrupted, which connects sites *i* and *m*

F^{max}: maximum number of fleets can be located

v_j: number of vehicles assigned to fleet *j*

Decision Variables:

F_j: binary variable deciding whether a fleet is located at fleet site *j*

y_{jm}: binary variable deciding whether site m is covered by fleet *j*

x_{ikj}: binary variable deciding whether, from fleet *j*, site *i* precedes site *k* (*i* < *k*) on a route

In the above terminology, fleet *j* denotes the fleet located at site *j*. We assume that $d_{ik} = d_{ki},\ t_{ik} = t_{ki},\ and\ \ p_{ik} = \ p_{ki},\ \forall i\ and\ \forall k \in M.$ Besides,$d_{ii} = \ t_{ii} = p_{ii} = \ 0, \forall i \in M, and$

y_{jj}=1, $\forall j \in M$. The two parameters, b_j and B_j, the minimum and the maximum number of sites each fleet will cover. These two limits decide the boundary of the number of sites for the fleet to perform its activity efficiently. To enhance the efficiency of the fleet's activity, each fleet should not cover too few or too many sites. The distance between a DRC and a site, d_{jm}, is also an essential factor of efficient operations since, generally speaking, the shorter the distance, the faster the delivery time, but it would not always be applied.

The total routing distance (*TRD*) and the total routing time (*TRT)* from all fleets to the sites are given by

$$TRD = \sum_{j\in M} F_j \left\{ \sum_{i\in M} \sum_{k\in M} y_{ik} x_{ikj}\, d_{ik} \right\}. \quad (1)$$

$$TRT = \sum_{j\in M} F_j \left\{ \sum_{i\in M} \sum_{k\in M} y_{ik} x_{ikj}\, t_{ik} \right\}. \quad (2)$$

Now, the probability of routing all given sites without being disrupted (*PRWD*) is given by

$$PRWD = \prod_{j\in M} F_j \prod_{i\in M} \prod_{k\in M} y_{ik}\, x_{ikj}(1 - p_{ik}). \quad (3)$$

Taking the natural log of both sides yields

$$\ell n\,(PRWD) = \sum_{j\in M} F_j \left\{ \sum_{i\in M} \sum_{k\in M} \ell n\,[\, y_{ik} x_{ikj}\,(1 - p_{ik})] \right\}. \quad (4)$$

Now, let *LPR* represent the left-hand side of Eq. (4), then

$$LPR = \sum_{j\in M} F_j \left\{ \sum_{i\in M} \sum_{k\in M} \ell n\,[y_{ik} x_{ikj}\,(1 - p_{ik})] \right\}. \quad (5)$$

Then, *PRWD* in (3) is rewritten as

$$PRWD = e^{LPR}. \quad (6)$$

Let the nonnegative deviation variables, (δ_1^+, δ_1^-), (δ_2^+, δ_2^-), and (δ_3^-, δ_3^+), denote the amounts by which each value of the three performance metrics deviates from the target values for the

three performance, TRD^*, TRT^*, *and* LPR^*. Then, the deviation variables are given by (see Ragsdale, 2018)

$$TRD\ in\ (1) + \delta_1^- - \delta_1^+ = TRD^*, \tag{7}$$

$$TRT\ in\ (2) + \delta_2^- - \delta_2^+ = TRT^*, \tag{8}$$

$$LPR\ in\ (5) + \delta_3^- - \delta_3^+ = LPR^*, \tag{9}$$

Let α_q^+ and α_q^- b relative importance weights attached to the overachievement and underachievement deviation variables. For analysis, we set the sum of all weights equal to one, that is, $\alpha_1^+ + \alpha_2^+ + \alpha_3^- = 1$, and each weight is a value between 0 and 1. Then, the weighted sum of the percentage deviations is defined as

$$Z(\alpha) = \alpha_1^+ \frac{(\delta_1^+ + \delta_1^-)}{TRD^*} + \alpha_2^+ \frac{(\delta_2^+ + \delta_2^-)}{TRT^*} + \alpha_3^- \frac{(\delta_3^+ + \delta_3^-)}{LPR^*}. \tag{10}$$

Setting up Equation (10) as an objective function, we formulate the FLAR design problem as a mixed-integer quadratic programming (MIQP) model shown below:

Minimize $Z(\alpha)$ in (10)

subject to

$$\sum_{j \in M} y_{jm} = 1, \qquad \forall m \in M \tag{11}$$

$$\sum_{j \in M} F_j = F^{max}, \tag{12}$$

$$y_{jm} \le F_j, \qquad \forall j\ and\ \forall m \in M \tag{13}$$

$$F_j \cdot b_j \le \sum_{m \in M} y_{jm} \le F_j \cdot B_j, \qquad \forall j \in M \tag{14}$$

$$x_{ikj} \le y_{ji}, \qquad \forall j \in M, \forall i \in M, and\ \forall k \in M \tag{15}$$

$$\sum_{i=1}^{k-1} x_{ikj} + \sum_{m=k+1}^{M} x_{kmj} = 2y_{jk}, \quad \forall j \in M \ \ and \ \forall k \neq j \tag{16}$$

$$\sum_{i=1}^{M} x_{ijj} + \sum_{m=1}^{M} x_{jmj} = 2v_j F_j, \quad \forall j \in M \tag{17}$$

$$\sum_{i,k \in S} x_{ikj} \leq (|S| - 1) y_{jk}, \ \ S \subset M, 3 \leq |S| \leq |M| - 3, \ \ \forall j \in M \tag{18}$$

Equations (7)-(9).

Constraints (11) make certain that a DRC covers each affected site. Constraints (12) define the maximum number of facilities to be built. Constraints (13) ensure that a selected DRC can only cover each site. Constraints (14) make sure that the fleet j must cover at most B_j and at least b_j sites. Constraints (15) make sure that a route should be formed at the location of fleet j. Constraints (16) ensure that each site in a route starting from fleet j should be accessible from two adjacent sites. Constraints (17) make sure that there should have $2v_j$ edges active at the facility j. Constraints (18) are called the subtour elimination constraint, where *S* denotes a subtour, which is a route of a set of the affected sites and not visiting all sites a vehicle is supposed to visit.

Suppose F^{max} is greater than or equal to 2, and each fleet can start its mission at the same time; then, the mission will be completed by the longest time among the routing time from each facility. The mission completion time (*MCT*) is given by

$$MCT = \max_{\forall j} \{ F_j \left\{ \sum_{i \in M} \sum_{k \in M} y_{ik} x_{ikj} \, t_{ik} \right\}, \forall F_j = 1. \tag{19}$$

Rather than *TRT* in (2), *MCT* in (19) can be applied in the formulation of the GP model.

DATA ENVELOPMENT ANALYSIS-BASED METHODS

Classical DEA

The mathematical model of C-DEA, which is called a multiplier DEA (*m*-DEA) model, may be stated as (see Zhu, 2014)

Objective Function: Maximize the efficiency rating θ for DMU_j

$$Max\ \theta_j = \frac{\sum_{r=1}^{s} u_r O_{rk}}{\sum_{i=1}^{m} v_i I_{ik}}. \quad (20)$$

This is subject to the constraint that when the same set of *u* and *v* coefficients (or weights) is applied to all other DMUs being compared, no DMU will be more than 100% efficient as follows:

$$DMU_j, \frac{\sum_{r=1}^{s} u_r O_{rj}}{\sum_{i=1}^{m} v_i I_{ij}} \leq 1, \forall j = 1, 2, \dots N \quad (21)$$

$$u_1, \dots, u_s > 0\ and\ v_1, \dots, v_m \geq 0,$$

where

N = number of DMUs under evaluation in the DEA analysis

θ = efficiency rating of the DMU_k being evaluated by DEA

O_{rj} = amount of output r generated by DMU_j, j =*1, 2, …, N*

I_{ij} = amount of input i used by DMU_j

i = index of inputs used by the DMUs, *i = 1, 2, …, m*

r = index of outputs generated by the DMUs, *r= 1, 2, …, s*

u_r = weight or coefficient assigned by DEA to output r

v_i = weight or coefficient assigned by DEA to input i

Now, the above model given by (20)-(21) can be transformed to the following linear programming (LP) problem, which is called a CRS (Constant Returns to Scale) *m*-DEA model:

$$Max\ ES_k = \sum_{r=1}^{s} u_r O_{rj} \quad (22)$$

subject to

$$\sum_{i=1}^{m} v_i I_{ik} = 1, \quad (23)$$

$$\sum_{r=1}^{s} u_r O_{rj} - \sum_{i=1}^{m} v_i I_{ij} \leq 0, j = 1, \dots, N, \quad (24)$$

$$u_r, v_i \geq 0, r = 1 \dots, s; i = 1 \dots, m.$$

ES_k^* denotes the optimal value of the objective function in (22) corresponding to the optimal solution (u^*, v^*), and DMU_k is defined to be efficient if $ES_k^* = 1$. DEA models can be either output-oriented or input-oriented, depending upon the rationale for conducting DEA. The model given by (22)-(24) is called an input-oriented CCR model, and ES_k^* is called CRS efficient score (ES).

Cross Efficiency DEA

The most popular DEA-based ranking method is cross-efficiency (CE) DEA. The CE-DEA method, which consists of two phases, was proposed to rank DMUs with the central idea of using DEA to do peer evaluation rather than pure self-evaluation (see Sexton et al., 1986). The weights or multipliers from the first phase are applied to all DMUs to get the cross-efficiency score (CES) for each DMU in the second phase. In the first phase, let E_{jj} represent the efficiency score for DMU_j, which is obtained by solving the following LP model:

$$max \quad E_{jj} = \sum_{r=1}^{s} u_{rj} O_{rj}, \tag{25}$$

subject to

$$\sum_{i=1}^{m} v_{ij} I_{ij} = 1, \tag{26}$$

$$\sum_{r=1}^{s} u_{rw} O_{rw} - \sum_{i=1}^{m} v_{ij} I_{iw} \leq 0, w = 1, \dots, n, \tag{27}$$

$$u_{rj}, v_{ij} \geq 0, r = 1, \dots, s; i = 1, \dots, m$$

To denote the peer evaluation, let E_{jw} represent the DEA score for the rated DMU_w, w = 1, 2, …, n, using the optimal weights /multipliers that a rating DMU_j has chosen in the model (25)-(27). Now, E_{jw} is given by

$$E_{jw} = \frac{\sum_{r=1}^{s} u_{rj}^* O_{rw}}{\sum_{i=1}^{m} v_{ij}^* I}, \qquad j \; and \; w = 1, \dots, n. \tag{28}$$

Note that E_{jj}, the DEA score for the self-evaluation of Phase I, is the same as E_j and can be obtained using the model by (25)-(27). Then, the CE score for DMU_w is defined as follows:

$$CE_w = \frac{1}{n}\sum_{j=1}^{n} E_{jw} \tag{29}$$

For the application of CE-DEA, see Gavgani and Zohrehbandian (2014), Hong and Jeong (2017), Hou et al. (2018), Hong and Jeong (2020), and Hong (2022).

Super efficiency DEA

The super-efficiency score (SES) is obtained from the C-DEA model after a DMU under evaluation is excluded from the reference set of the C-DEA models (see Anderson and Peterson, 1993). The resulting model is called a SE-DEA model with significance for discriminating among efficient DMUs. The SES for DMU_k is expressed as

$$max\ SES_k = \sum_{r=1}^{s} u_r y_{rk}, \tag{30}$$

subject to

$$\sum_{i=1}^{m} v_i x_{ik} = 1, \tag{31}$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, j \neq k, \tag{32}$$

$$u_r, v_i \geq 0, r = 1 \ldots, s; i = 1 \ldots, m.$$

In the above SE model given in (30)-(32), efficient DMUs are not compared to the same standard since the frontier constructed from the remaining DMUs changes for each efficient DMU to be rated. Consequently, the SESs of efficient DMUs are allowed to have higher values than 1. See Nayebi and Lotfi (2016) and Deng et al. (2018) for the application of SE-DEA.

This study utilizes CES and SES for evaluating and ranking the generated FLAR. We propose the following procedure for solving the GP model to generate FLAR schemes and evaluating those generated schemes using the DEA methods with two inputs, *TRD* and *TRT* (or *MCT*), and one output, *PRWD*.

Procedure

Step 1: [**Setup GP model and solve the model to generate FLAR schemes]**

(i) Formulate the GP model and set various values of the weight set, $\boldsymbol{\alpha} = \{\alpha_1^+, \alpha_2^+, \alpha_3^-\}$, where each weight varies between 0 and 1 with an increment of Δ, where Δ =1/d, d is an integer, and Δ should be a finite decimal.

(ii) For each set of weights, solve the GP model in (7)-(18) to generate DMU_κ, κ=1, 2, …, Ω, which is the optimal solution for the given weight set.

Step 2: [**Applying CE and SE -DEA for evaluating the schemes generated in Step 1]**

(iii) Compute ES in (22) for each DMU with one output, *PRWD*, and two inputs, *TRD* and *TRT/MCT*, generated in (ii) of Step 1.

(iv) Obtain the CE matrix by solving the CE models for the generated DMUs

(v) Compute CES in (29) and SES in (30) for each DMU for all weight sets.

(vi) Find the averages of CES, $\overline{CE}$. Rank DMUs based on the efficiency scores, ES, CE, and SES values computed in (iii) and (v).

NUMERICAL EXAMPLE

We consider a numerical example with nine sites that Goetschalckx (2011) illustrates, shown in Figure 2. The mission is to locate a specified number of fleets; each fleet covers certain sites, so all sites are covered. Then, each fleet should visit the assigned sites. Tables 1, 2, and 3 list hypothetically pre-determined input data, routing distance, routing time, and the disruption probability used for the numerical example, and we set the number of vehicles in each fleet, v_j, equal to 1. First, setting the maximum number of fleets equal to 1, F^{max}=1, and setting each goal as a single objective value, we find the optimal value of each performance metric (PM), as shown in Table 4. We observe that two extreme weight sets, $\boldsymbol{\alpha} = \{1, 0, 0\}$ and {0, 1, 0}, generates the same values of three PMs, and *PRWD* is just above 0.5 for these two weight settings. Now with the weight set, $\boldsymbol{\alpha} = \{0, 0, 1\}$, the *PRWD* jumps to 0.8885, but the other two PMs, *TRD* and *TRT*, are sacrificed with much larger values than the first two cases, $\boldsymbol{\alpha} = \{1, 0, 0\}$ and {0, 1, 0}. We find FLAR schemes for these two results, and Figure 3 depicts these two FLAR network schemes, supporting the results in Table 4. As expected, the FLAR scheme with $\boldsymbol{\alpha} = \{0, 0, 1\}$ shows a more complicated route sequence to maximize the *PRWD* only, ignoring the other two PMs, *TRD* and *TRT*. With F^{max}=1, the fleet can be located at any site, generating the same values of three PMs. The decision-maker can select the location based on many strategic factors.

Table 1. Input data of routing distance used for the numerical example

	C1	C2	C3	C4	C5	C6	C7	C8	C9
C1	0	22.06	42.56	41.38	17.88	24.62	41.62	40.4	29.22
C2	22.06	0	32.92	∞	29.12	39.8	43.64	∞	48.34
C3	42.56	32.92	0	43.96	32	66.76	22.1	65.1	52
C4	41.38	∞	43.96	0	25.28	∞	23.38	28.14	20.64
C5	17.88	29.12	32	25.28	0	41.1	24.52	∞	21.64
C6	24.62	39.8	66.76	∞	41.1	0	∞	47.32	42.08
C7	41.62	43.64	22.1	23.38	24.52	∞	0	48.78	37.26
C8	40.4	∞	65.1	28.14	∞	47.32	48.78	0	13.44
C9	29.22	48.34	52	20.64	21.64	42.08	37.26	13.44	0

Table 2. Input data of routing time used for the numerical example

	C1	C2	C3	C4	C5	C6	C7	C8	C9
C1	0	42.1	30.5	29.5	20	30.5	35.8	40.4	30.4
C2	42.1	0	25.7	∞	19.3	40.5	41.5	∞	35.8
C3	30.5	25.7	0	40.5	42.8	50.8	15.5	35.8	45.5
C4	29.5	∞	40.5	0	30.6	∞	15.8	25.4	18.2
C5	20	19.3	42.8	30.6	0	28.9	25	∞	20.5
C6	30.5	40.5	50.8	∞	28.9	0	∞	35.3	40.5
C7	35.8	41.5	15.5	15.8	25	∞	0	29.9	35.2
C8	40.4	∞	35.8	25.4	∞	35.3	29.9	0	10.4
C9	30.25	35.8	45.5	18.2	20.5	40.5	35.2	10.4	0

Table 3. Input data of disruption probability used for the numerical example

	C1	C2	C3	C4	C5	C6	C7	C8	C9
C1	0	0.03	0.125	0.014	0.175	0.02	0.2	0.035	0.08
C2	0.035	0	0.02	0.9999	0.0012	0.18	0.05	0.9999	0.075
C3	0.01	0.058	0	0.08	0.045	0.15	0.05	0.06	0.01
C4	0.014	0.9999	0.1	0	0.02	0.9999	0.006	0.002	0.0337
C5	0.0025	0.075	0.085	0.09	0	0.12	0.04	0.9999	0.02
C6	0.1	0.06	0.008	0.9999	0.12	0	0.9999	0.07	0.014
C7	0.2	0.02	0.15	0.02	0.01	0.9999	0	0.036	0.098
C8	0.002	0.9999	0.024	0.25	0.9999	0.003	0.05	0	0.02
C9	0.125	0.036	0.084	0.18	0.19	0.085	0.095	0.065	0

Figure 2. Site locations and routes

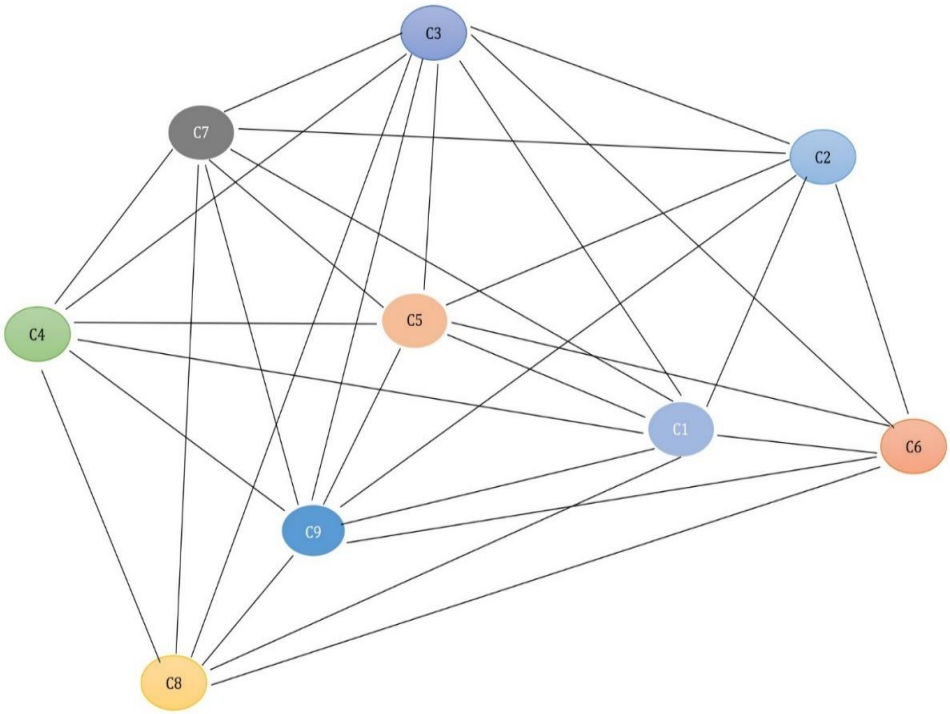


Table 4. Three extreme cases with F^{max}=1

$\boldsymbol{\alpha} = \{\alpha_1^+, \alpha_2^+, \alpha_3^-\}$	*TRD* (miles)	*TRT* (min.)	*PRWD*
(1, 0, 0)	231.4	190.7	0.5644
(0, 1, 0)	231.4	190.7	0.5644
(0, 0, 1)	374.9	281.9	0.8885

Now, setting F^{max}=2, the GP model with five goals is formulated using Eq. (11)-(23) and solved for various weight values by setting an increment value, Δ, equal to 0.1 after several Δ values are tried. To decide the value of Δ, we follow (i) in Step 1. With Δ < 0.2, the number of computations increases, and solving the model yields many identical FLAR configurations. In contrast, with Δ > 0.2, the number of the generated configurations tends to be small. It implies that several efficient configurations could be missing. A total of 21 configurations with Δ = 0.2 arise out of the combinations of the setting of **α** under the constraint, $\alpha_1^+ + \alpha_2^+ + \alpha_3^- = 1$. Each configuration stands for a FLAR network scheme, which is considered a DMU for the DEA model. After solving the GP model, some network configurations yield the same PM values for all three. After combining the same designs, these network schemes are consolidated and reduced to eight (8) DMUs. Table 5 shows the eight FLAR schemes and the corresponding weight set, the values of three PMs using the total routing time (*TRT*), ES, CES, SES, and the corresponding rankings based on the efficiency scores.

As shown in Table 5, all three efficiency scores rank DMU_{14} as #1, which the GP model generates with the weight set **α** = {0.4, 0.4, 0.2}. The #2 ranked DMU, DMU_5, with **α** = {0.0, 0.8, 0.2}, indicates that *TRD* and *TRT* are somehow interrelated. The correlation between *TRD* and *TRT* turns out to be 0.70. Note that DMUs with $\alpha_3^- = 0$, DMU_6, DMU_{11}, DMU_{15}, DMU_{18}, and DMU_{21}, rank the lower parts, #4, #5, #6, #7, and #8, as the resulting values of the *PRWD* are very low compared to other cases. We observe that the network schemes will be ranked low, ignoring the only output of the probability of not being disrupted. From these results, considering the risk of route disruptions in the design of FLAR would play an essential factor in determining the efficiency of the network scheme.

As described before, if the two fleets start their mission simultaneously, the mission completion time, *MCT* in (39), rather than *TRT*, can be an input for the DEA method to be applied. Table 6 shows the eight FLAR schemes using *MCT* as an input, its corresponding ES,

CES, and SES, and the rankings based on these efficiency scores. The rankings of DMUs are not all identical when *TRT* is used as an input. Now, DMU_5, with a shorter *MCT* and a higher *PRWD* than DMU_{14}, emerges as a top-ranked DMU, as ES and CES indicate, while DMU_{14} has the highest rank in Table 5, is still ranked the top for ES and SES. Figure 5 displays FLAR network schemes with rankings of #1 through #4, while Figure 6 shows those with rankings of #5 through #8.

All the results shown in Tables 5 and 6 and the FLAR network schemes depicted in Figures 4 and 5 would be helpful for the decision-maker to decide which FLAR scheme to adopt or implement based on the mission characteristics, preferences, and/or the current status of the route.

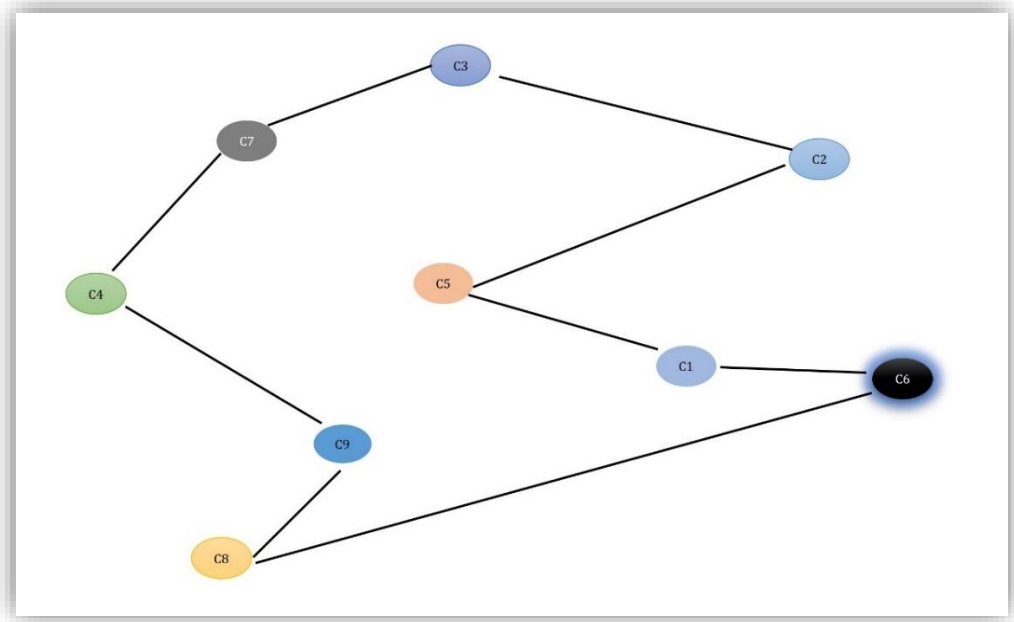
SUMMARY AND CONCLUSIONS

This paper proposes a procedure of designing FLAR network schemes under the risk of route disruptions, which is missing from the existing literature on FLAR modeling problems. The strategic design of the efficient FLAR network scheme could be an important task for completing missions successfully. But, in reality, we are experiencing many cases of route disruptions. Each route is subject to disruptions caused by various sources or reasons, including natural disasters such as climate-related disasters. These disruptions are usually difficult to predict and are sometimes inevitable. This study considers three performance measures, the total routing distance (*TRD)*, the total routing time (*TRT)*, and the probability of routing without being disrupted (*PRWD*), which would be directly related to completing the mission.

This paper formulates the FLAR network design problem with the risk of route disruptions through the two-step framework to identify efficient FLAR network schemes. In the first step, a GP model is formulated for the FLAR problem to find the optimal location of the fleet, allocation, and routing decisions. Solving the GP model generates an optimal FLAR network scheme for a given weight assigned to each performance metric. Consequently, solving the GP model for various values of the weights would provide many network alternatives. Treating each alternative as a DMU with *PRWD* as output and *TRD* and *TRT/MCT* as two inputs, three DEA techniques, classical DEA, cross-efficiency, and super-efficiency methods are combined and applied to find the efficient and the most efficient FLAR network schemes. These efficient FLAR network schemes would help the decision-makers select the final one.

Figure 3. FLAR network schemes with F^{max} =1

(i) $\alpha = \{1, 0, 0\}$ **or {0, 1, 0}**



(ii) $\alpha = \{0, 0, 1\}$

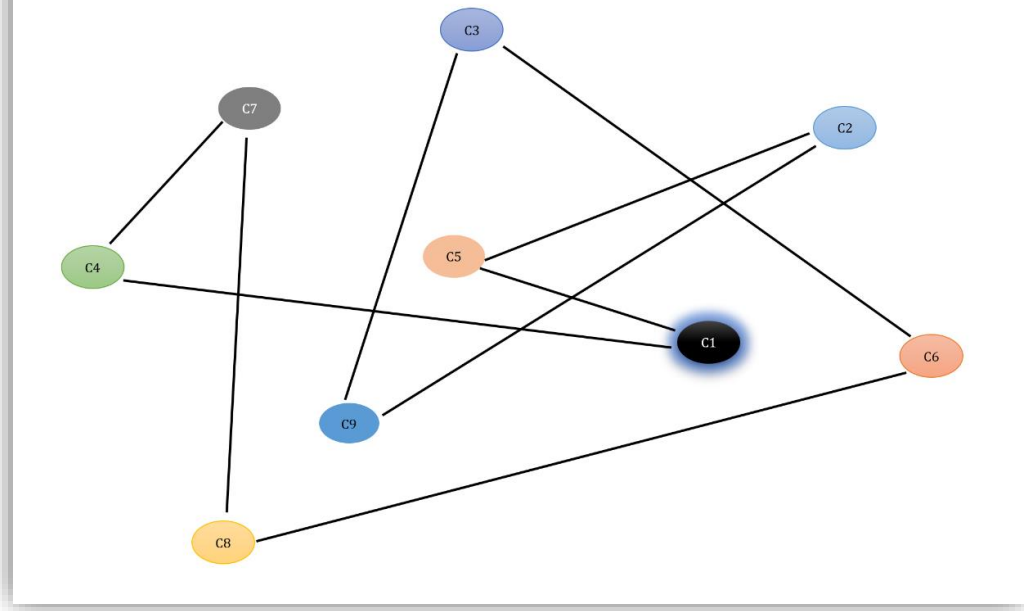


Table 5. Performance metrics using *TRT*, efficiency scores, and rankings for FLAR schemes with F^{max}=2

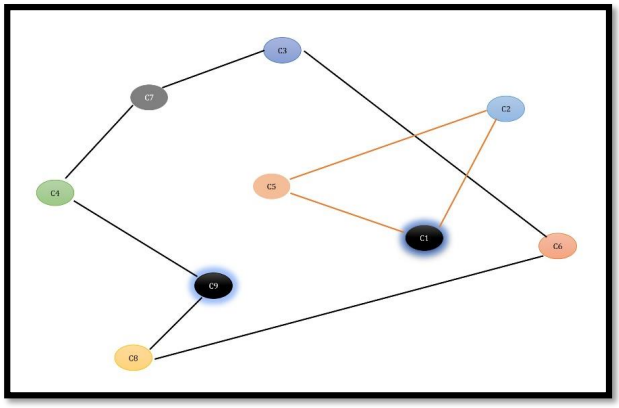
No	DMU #	α=(α1, α2, α3)	*TRD* (miles)	*TRT* (min.)	*PRWD*	ES [R]	CES [R]	SES [R]
			Input	Input	Output			
1	1	(0.0, 0.0, 1.0)	334.9	268.1	0.8760	0.9243 [3]	0.9069 [3]	0.9243 [3]
2	5	(0.0, 0.8, 0.2)	300.9	243.5	0.8526	0.9905 [2]	0.9743 [2]	0.9905 [2]
3	6	(0.0, 1.0, 0.0)	251.9	208.0	0.4441	0.6040 [6]	0.5970 [6]	0.6040 [6]
4	11	(0.2, 0.8, 0.0)	256.1	209.4	0.5544	0.7489 [5]	0.7386 [5]	0.7489 [5]
5	14	(0.4, 0.4, 0.2)	262.7	227.4	0.8039	1.0000 [1]	1.0000 [1]	1.0799 [1]
6	15	(0.4, 0.6, 0.0)	253.7	209.7	0.3433	0.4631 [8]	0.4579 [8]	0.4631 [8]
7	18	(0.6, 0.4, 0.0)	247.4	213.3	0.6463	0.8569 [4]	0.8561 [4]	0.8569 [4]
8	21	(1.0, 0.0, 0.0)	227.2	243.5	0.3969	0.5709 [7]	0.4885 [7]	0.5709 [7]

Table 6. Performance metrics using *MCT*, efficiency scores, and rankings for FLAR schemes F^{max}=2

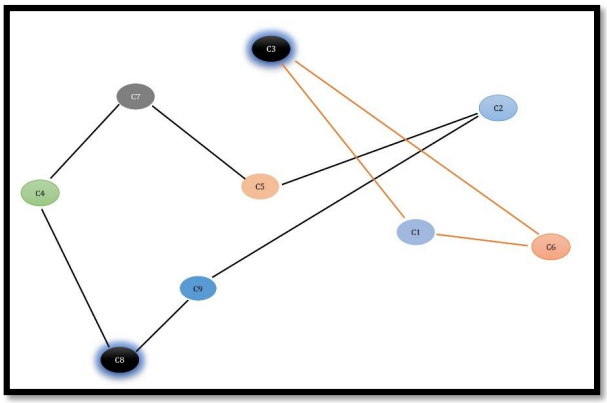
No	DMU #	α=(α1, α2, α3)	*TRD* (miles)	*MCT*(min.)	*PRWD*	ES [R]	CES [R]	SES [R]
			Input	Input	Output			
1	1	(0.0, 0.0, 1.0)	334.9	142.0	0.8760	0.9529 [3]	0.9081 [3]	0.9229 [4]
2	5	(0.0, 0.8, 0.2)	300.9	131.7	0.8526	1.0000 [1]	0.9723 [1]	1.0645 [2]
3	6	(0.0, 1.0, 0.0)	251.9	128.6	0.4441	0.5929 [6]	0.5717 [6]	0.5929 [6]
4	11	(0.2, 0.8, 0.0)	256.1	123.9	0.5544	0.7409 [5]	0.7159 [5]	0.7409 [5]
5	14	(0.4, 0.4, 0.2)	262.7	146.0	0.8039	1.0000 [1]	0.9626 [2]	1.0799 [1]
6	15	(0.4, 0.6, 0.0)	253.7	155.7	0.3433	0.4422 [8]	0.4109 [8]	0.4422 [8]
7	18	(0.6, 0.4, 0.0)	247.4	106.9	0.6463	0.9339 [4]	0.9006 [4]	0.9339 [3]
8	21	(1.0, 0.0, 0.0)	227.2	130.4	0.3969	0.5708 [7]	0.5432 [7]	0.5708 [7]

Figure 4. Efficient FLAR network schemes with F^{max}=2

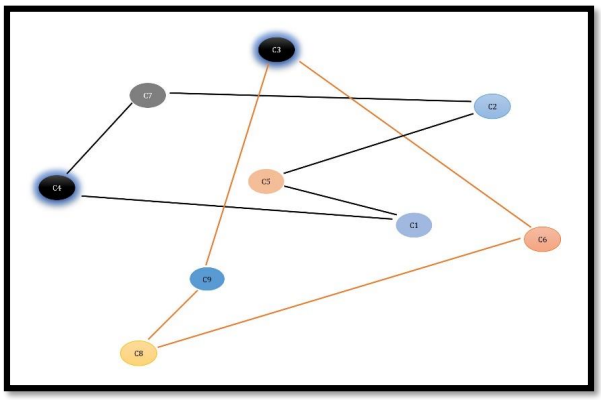
DMU$_{14}$



DMU$_{5}$



DMU$_{1}$



DMU$_{18}$

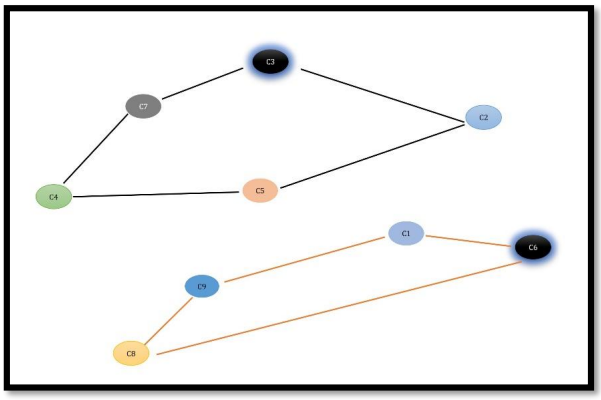
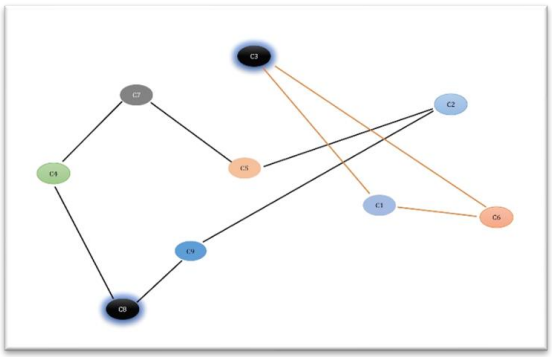
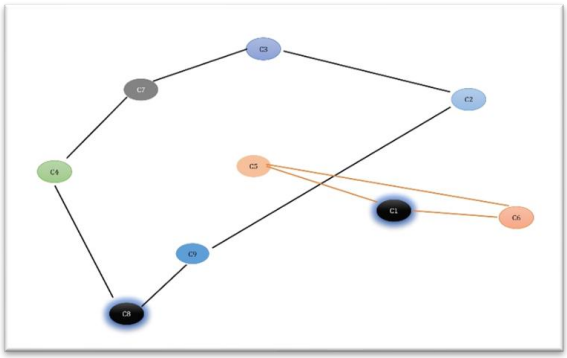


Figure 5. Inefficient FLAR network schemes F^{max}=2

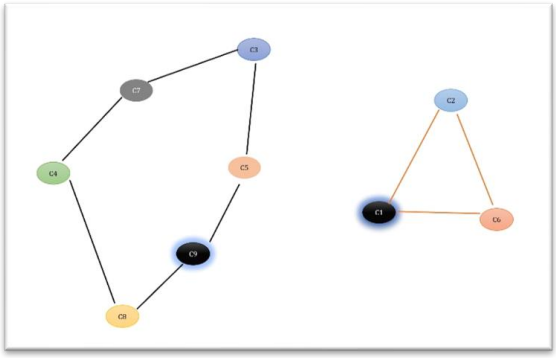
DMU$_{11}$



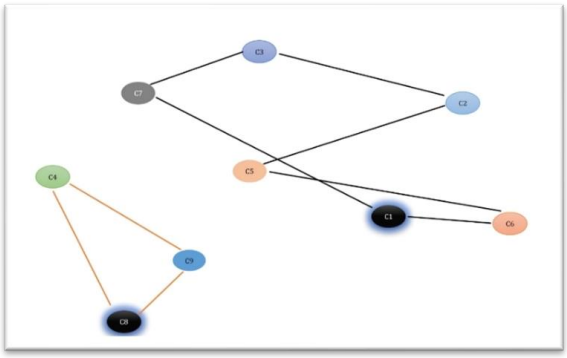
DMU$_{6}$



DMU$_{21}$



DMU$_{15}$



The contribution of this study is to propose the framework for the FLAR network design problem. Each alternative generated by solving the GP model yields the optimal value of outputs and inputs, so DEA methods can be applied to measure each network alternative's relative efficiency and identify the most efficient network scheme(s). The peer-evaluation-based cross-efficiency DEA (CE-DEA) method was proposed to overcome the critical weakness of the C-DEA based on self-evaluation. As mentioned before, the critical drawback of the CE-DEA method is that the cross-efficiency score (CES) generated by CE-DEA depends upon the kinds of optimization software. In addition, the effect of self-evaluation/peer-evaluation proportion on the CES has not been cleared. The super-efficiency DEA (SE-DEA), which does not have such an issue of inconsistent CES of CE-DEA, uses the two adjacent DMUs in terms of efficiency to generate an SE score. This study applies a CE-based SE-DEA approach to identify the most efficient DRCLAR network scheme(s). At the same time, we consider the effect of self-evaluation/peer-evaluation proportion, which is the other weakness of CE-DEA, on the rankings of DMUs.

The proposed approach to the FLAR design problem would provide many insights to practitioners as well as researchers. Using a case study, we observe that the proposed framework of combining the GP model and three DEA methods works very well regarding identifying and ranking efficient schemes. Some goals can be transformed into constraints, depending on the possible situation of the mission or the decision-maker’s requirements on the performance measures. As mentioned before, it would be problematic to change fleet locations, but actual fleet allocation and routing schemes could be modified according to decision-maker’s discretion. Besides, the proposed framework could be applied to design various fleet strategic and operational network systems if their performance measures can be classified into inputs or outputs.

This study assumes that its allocated sites are not covered if a route is disrupted. For future research, it would enhance this research if the concept of backup routes for the case of route shutdown due to disruptions is considered. Thus, an emergency backup routing plan should be made if some routes are disrupted. That is how to mitigate the impact of disruptions on the general supply chain network system. This study assumes that only DRCs are subject to disruptions. Besides, the maximum allowable routing times to any site would be another critical factor for completing a mission. It will surely enhance this research if the constraint for the maximum delivery times is also considered.

REFERENCES

Anderson, A., & Petersen, C. N. (1993). A procedure for ranking efficient units in DEA. *Management Science*, 39(10), 1261-1264.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2(6), 429-444.

Charnes, A., Haag, S., Jaska, P., & Semple, J. (1992). Sensitivity of efficiency classifications in the additive model of data envelopment analysis. *International Journal of Systems Science*, 23(5), 789-798.

Daskin, M.S. (2013). *Network and discrete location: Models, algorithms, and applications.* 2nd Edition, Wiley, New York.

Deng, Z., Gao Y., Liang, B., & Morrison, A. M. (2018). Efficiency evaluation of hotel operations in Mainland China based on the super-efficiency SBM model. *Tourism Economics*, 26(2), 276-298.

Doyle, J., & Green, R. (1994). Efficiency and cross-efficiency in DEA: Derivations, meanings, and uses. *Journal of Operational Research Society*, 45(5), 567-578.

Fang, L., & Li, H. C. (2015). Multi-criteria decision analysis for efficient location-allocation problem combining DEA and goal programming. RAIRO-Operations Research, 49(4), 753-772.

Farahani, R. Z., SteadieSeifi, M., & Asgari, N. (2010). Multiple criteria facility location problems: A survey. *Applied Mathematical Modelling*, 34 (7), 1689-1709.

Farahani, R. Z., Bajgan, H.R., Fahimnia, B., & Kaviani, M. (2015). Location-inventory problem in supply chains: A modeling review. *International Journal of Production Research*, 53(12), 3769-3788.

Gavgani, S.S., & Zohrehbandian, M. (2014). A cross-efficiency-based ranking method for finding the most efficient DMU. *Mathematical Problems in Engineering*, 2014, Article ID. 269768.

Goetschalckx, M. (2011). *Supply chain Engineering*, Spring, New York.

Hong, J.D., & Jeong, K. Y. (2017). Cross-efficiency based heuristics to rank decision-making units in data envelopment analysis. *Computers & Industrial Engineering*, 111, 320-330.

Hong, J.D., & Jeong, K. Y. (2019). Combining data envelopment analysis and multi-objective model for the efficient facility location-allocation decision, *Journal of Industrial Engineering International*, 15, 315-331.

Hong, J.D., & Jeong, K. Y. (2020). Cross-evaluation based super-efficiency DEA approach to designing disaster recovery center location-routing network schemes. *Journal of Humanitarian Logistics and Supply Chain Management*, 10(4), 485-508.

Hong, J. (2022). Applying Cross Efficiency Evaluation Methods for Multi-Objective Emergency Relief Supply Chain Network Model. *International Journal of Industrial and Systems Engineering*, 41(1), 19-40.

Hou, Q., Wang, M., & Zhou, X. (2018). Improved DEA cross-efficiency evaluation method based on ideal and anti-ideal points. *Discrete Dynamics in Nature and Science*, 2018, Article ID. 1604298.

Klimberg, R. K., & Ratick, S. J. (2006). Modeling data envelopment analysis (DEA) efficient location/allocation decision. *Computers & Operations Research*, 35(2), 457-474.

Lee, Y. C. (2019). Ranking DMUs by combining cross-efficiency scores based on Shannon's entropy. *Entropy*, 21(5), 467.

Liu, H.H., Song, Y.Y., & Yang, G. L. (2019). Cross-efficiency evaluation in data envelopment analysis based on prospect theory. *European Journal of Operational Research*, 273(1), 364-375.

Meza, D., & Jeong, K.Y. (2013). Measuring efficiency of lean sigma project implementation using data envelopment analysis at NASA. *Journal of Industrial Engineering and Management*, 6(2), 401-422.

Nayebi, M., & Lotfi F.H. (2016). The methods of ranking-based super efficiency. *International Journal of Research in Industrial Engineering*, 5(4), 43-66.

Panwar, A., Olfati, M., Pant M., & Snalsel, V. (2022). A review on the 40 years of existence of data envelopment analysis models: Historic development and current trends. *Archives of Computational Methods in Engineering*, 29, 5397-5426.

Paryzad, B., Najafi, E., Kazemipoor, H., & Pour, N.S. (2018). The new ranking method of the decision-making units in DEA: with an approach to modifying cross-efficiency method. *International Journal of Industrial and Systems Engineering*, 30(3), 387-400.

Ramanathan, R. (2006). Data envelopment analysis for weight derivation and aggregation in the analytic hierarchy process. *Computers & Operations Research*, 33(5), 1289-1307.

Sexton, T.R., Silkman, R.H., & Hogan, A.J. (1986). Data envelopment analysis: Critique and extensions, in Silkman, R. (Ed.): *Measuring efficiency: An assessment of data envelopment analysis*, 73-105, Jossey-Bass, USA.

Zhu, J. (2014). *Quantitative models for performance evaluation and benchmarking: data envelopment analysis with spreadsheets*. 3rd Edition, Springer, New York.

Zhu, J. (2022). DEA under big data: data-enabled analytics and network data envelopment analysis, *Annals of Operations Research*, 309, 761-783.

DECISION SCIENCES INSTITUTE
Improving Police Response with a Crisis Intervention Team

ABSTRACT

This research points to the deficits in the current law enforcement response to individuals experiencing mental health crises. Unfortunately, people in crisis are often jailed or hospitalized, due to a lack of resources and underutilized mental health facilities. Current statistics show that many people in jail have mental illness, but the treatment for mental illness is not accessible to those individuals, and hospitals have a limited number of beds. Simulation-based analyses of current police response procedures are presented in this paper to find improvements and propose changes based on the need for the patient, which ultimately will reduce the amount of individuals erroneously being jailed and increase the number of people getting right treatment.

KEYWORDS: Mental health crisis, Crisis intervention training, Law enforcement response, Discrete event simulation, Emergency response procedure

INTRODUCTION

Today police reform has garnered national attention due to events involving police violence, political unrest, and riots. There are many issues plaguing the relationship between police and communities, however ensuring that police have access to resources, training, and funding will allow for better community policing to improve this relationship. One area of focus for improvement is officers' interactions with people in crisis or those with mental illness. The idea here is to improve current police response procedures to include in-house psychologists/counselors to streamline the resources and get the subject the help they need in a much more efficient fashion. Improving the process of police response can help the community, those with mental illness, and the police departments struggling with these issues.

LITERATURE REVIEW

First, we will analyze the problems of the police response and lack of resources as it stands today (as the Pre-Intervention Model). Then we can evaluate how this process can be improved with the proposed Crisis Intervention Team (Post-Intervention Model).

Problems with Current Crisis Response

Mental illness, suicide prevention, and mental health awareness is a topic of increasing importance. Especially since the COVID pandemic, we have seen mental health issues on the rise, causing many to develop depression or worsen depression and it is estimated to affect at least 1 in 10 people in America (Reinberg, 2022). These issues affect everyone from a national level to the regional level in the Midwest. In an attempt to address the issues that police officers face when responding to mental health calls, many states have added additional training requirements of their police officers. Many states have enacted laws to ensure that officers get Crisis Intervention Training, which "aims to reduce the risk of serious injury or death during an emergency interaction between persons with mental illness and police officers" and to get the

person resources rather than taking them to jail (Rogers et al., 2019). In some states, the legislation requires officers to take a certification course of 40 hours for Crisis Intervention that includes information on various mental illnesses and de-escalation techniques (ILCS, 2021). In an ideal world, a responding officer would have many more options for a mentally ill person other than jail, but in many communities, the appropriate resources are not available and in order to protect themselves and others, jail becomes a default option. According to the *Treatment Advocacy Center*, in "44 states, a jail or prison holds more mentally ill individuals than the largest remaining state psychiatric hospital" and although many counties in the U.S. have both a jail and psychiatric facilities, more seriously mentally ill individuals end up in prison (Caroll, 2016). This is even more concerning when you consider that the leading cause of death for incarcerated persons is suicide, and it is on the rise. The Bureau of Justice Statistics tracked deaths from 2001 to 2019 and found that suicides increased 85% in state prison, 61% in federal prison, and 13 % in local Jails (Bureau of Justice, 2021). According to *Prison Policy Initiative*, more than a third of deaths in jail are due to suicide, a person is 3 times more likely to die of suicide in jail, and of those that die from suicide in jail, more than half occur in the first 9 days of incarceration (Wang, 2021). It is important to note that being in jail could contribute to this, but mental health and suicide are a problem in society too. In 2020, the CDC (Centers for Disease Control) cited suicide as the twelfth leading cause of death in the United States, and death by suicide was nearly twice the rate of homicide (Suicide, 2022). These shocking numbers highlight the need of resources for people suffering with mental illness and even substance abuse. For police officers, interactions with people in crisis can escalate quickly without the proper training, but police are often the first or only responder to arrive at the call, so training is imperative. However, even if every officer in the nation received Crisis Intervention Training, there would still be a need for more resources to keep people out of jail who need treatment.

The Crisis Intervention Team

Interactions between the police and the public have shown that the process could be improved in order to better serve those in crisis or with mental illness. Today, "more police departments are teaming with mental health clinicians—including psychologists—out in the field or behind the scenes via crisis intervention training. When these groups collaborate well, people with mental illness in crisis can access mental health care more easily, police experience less trauma and stress, and clinicians have an opportunity to make an even bigger difference in the community" (Abramson, 2021). The breakdown in the current process is really the lack of mental health professionals responding to the incident and access to mental health professionals and facilities. If a police department could incorporate some mental health professionals into their response effort, it would reduce the time spent with various people evaluating the person and streamline the process of getting people the help that they need. Also, this aids in utilizing a resource in town other than the hospital or jail, which both have holding limits, such as in-patient facilities and behavioral health facilities that can help people long-term. Hiring an in-house psychologist to respond to crisis calls with an officer would reduce the time spent for the officer in trying to allocate these resources. The officer could then respond to maintain safety in the situation in case a person becomes combative, but the psychologist (and medical personnel if necessary) could evaluate and help decide next steps including possible outpatient treatment. This would result in less people going to the emergency room and lessen the number of people that police must take to jail because there is no other option. A University in the Midwest region has a pilot program in which they are exploring this concept. The department will hire social workers, referred to as 'crisis counselors' who are not officers and will respond to calls with officers (Herman, 2021). The Lieutenant at this University states "The idea is: if we have a

mental health professional that can evaluate that person on the scene, we will cut down, hopefully, a lot of our transports (Herman, 2021)", and they will be able to refer the person to resources that may not otherwise be readily available, which cuts down on the number of people going to the ER and jail. According to the article, *Building Mental Health into Emergency Response*, the benefit of this teamwork between officers and counselors is that "team members use trauma-informed, harm-reduction techniques to de-escalate crises and, if necessary, transport clients to outpatient care, reducing unnecessary emergency room visits and jail time" (Abramson, 2021). Streamlining the response process with a Crisis Intervention Team could greatly cut down on stress and risks to the officer and patient, and it would help the community be happier, safer, and hopefully get access to the help they need.

HYPOTHESIS

While recognizing that each call is different and cannot be generalized, the current process can be modeled by simulating what could potentially happen from the time a call comes into the police department until its resolution (resulting in release or transport to the hospital or jail) via the discrete event simulation package, AnyLogic. Knowing that this process can be time-consuming and generally inefficient for the subject/caller, the hope is to improve the process (model) to incorporate a psychologist into the response creating a Crisis Intervention Team (CIT). Based on this research, the authors' theory is that this should streamline the process for the person in crisis and get them the necessary resources without long wait times and expedite the overall contact with officers to come to a resolution more quickly. This should reduce the amount of time an individual waits to speak with a counselor, they can be treated quickly, and if they do need to go the ER or jail, the on-site counselor can help make those decisions quickly and safely with the officer.

PRE/POST-INTERVENTION MODELS

Base Model Summary Flow Chart (Pre-Intervention)

A mental health crisis usually starts as a 911 call, from someone having a mental health issue. These range from individuals who are feeling depressed, those who are under the influence of drugs or alcohol causing them mental distress, all the way to more serious calls of attempted suicide. An officer is sent to the patient and from there the officer must evaluate the situation and determine whether an ambulance (rescue) is needed or if jail is the more suitable option. It should be noted that police at a university may have more resources than municipal, county, or state police, because they often have a medical or counseling center for students. Such is the case at the university in the Midwest region, where the officer may utilize the resource of the counseling center during daytime hours, or after hours call an on-call counselor, but this usually takes a significant amount of time. Wait times could be 30 minutes to an hour or longer depending on their availability. The other options remain the same because if there is an injury they may need to go to the hospital, or if there was a crime, they would need to go to jail as shown in Figure 1.

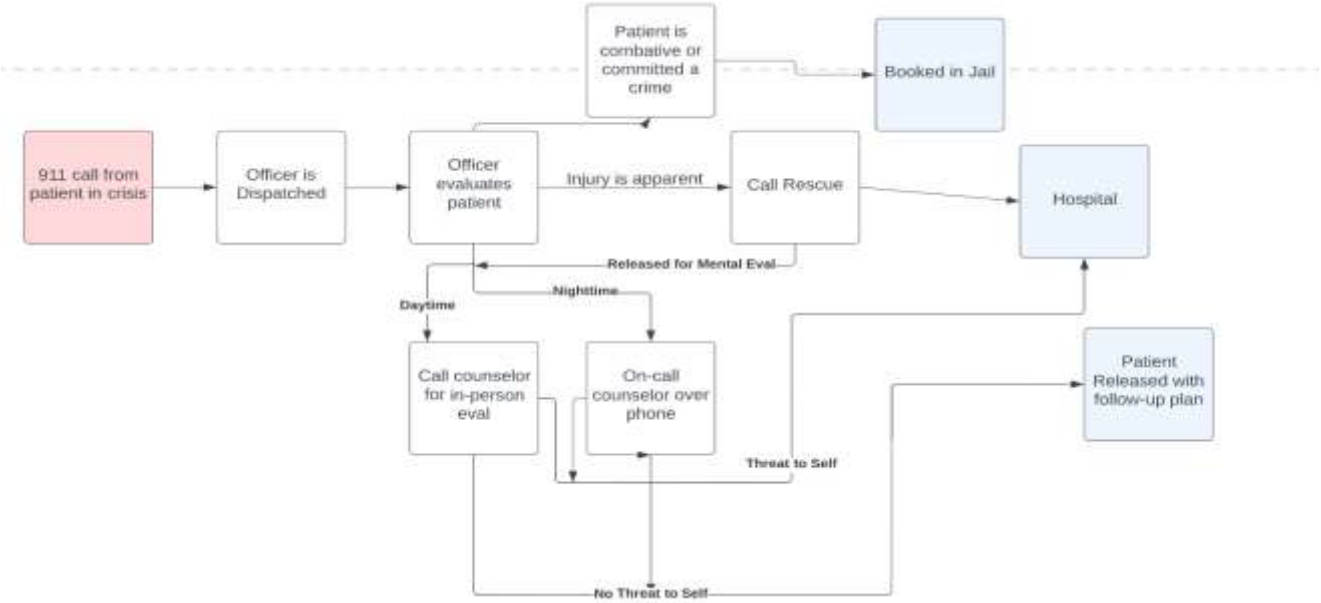


Figure 1. Pre-intervention model flowchart

Crisis Intervention Team Models (Post-Intervention)

All crisis calls start with a call to the police department or 911. The Crisis Intervention Team would then be dispatched, consisting of both the officer and psychologist. If there is a danger to safety or a need for first aid, the officer can take care of mitigating the threat and calling rescue while the psychologist does their mental health assessment. This allows for the person to be evaluated and treated quickly, with the psychologist making the recommendation for the next steps for the patient's care and having access to more resources at the scene should cut down on those who end up in jail. An in-house psychologist could also mean the ability to refer patients to places like in-patient mental health treatment facilities that further take the burden off the hospitals and jails. This is represented below in Figure 2:

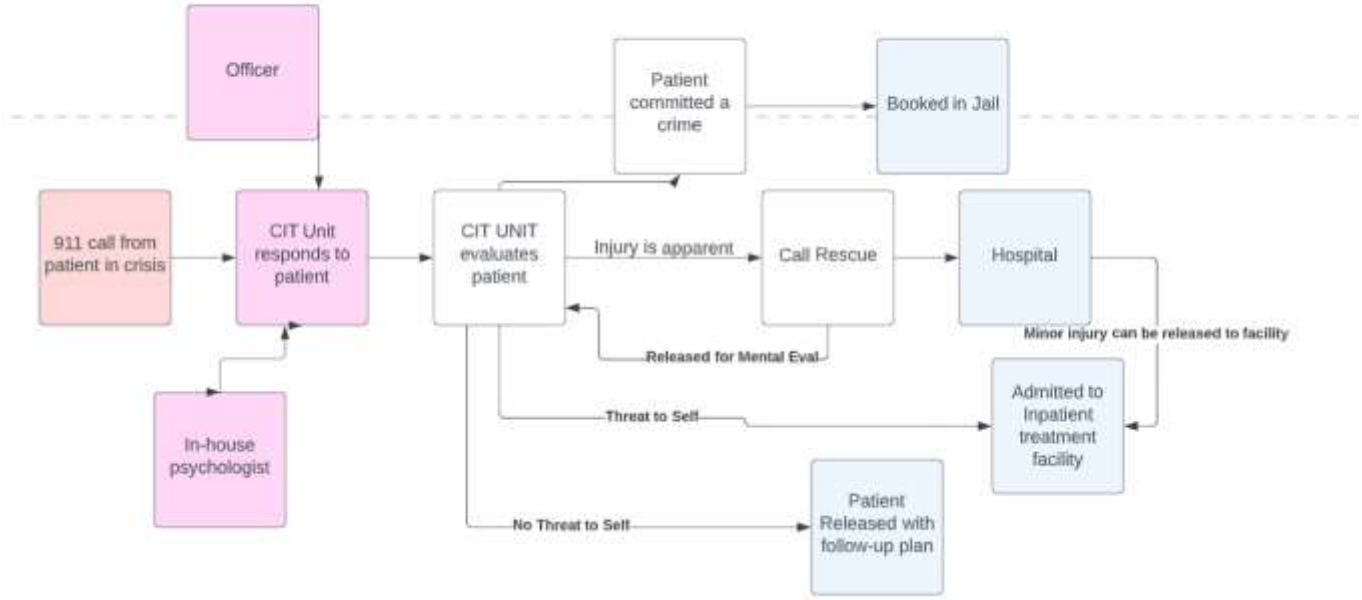


Figure 2. Post-intervention model flowchart

SIMULATION

The methods used to simulate these real-world scenarios are listed below using Any Logic. Considering that there is no real-world model to compare this to, the authors chose the data points that would best represent improvements in the model and most importantly for the patient agent.

Pre-Intervention Model (Base Model)

The simulation for the pre-intervention model in Figure 3 shows how the call progresses from the initial crisis, to receiving resources/treatment, to the resolution. Individuals are moved through the process to simulate real life, and will have a statistical probability of being released, booked in jail, or taken to the hospital. This simulation model is built on the idea that there is the availability of a counseling center/after hours on-call counselor (as there may be a university in the Midwest region) for individuals to utilize when available. In general, a counseling center has a disadvantage in that it could have long wait times, and it is only open during business hours, and after-hours it is only good if someone is available. This is not always practical if the person is injured or is uncooperative.

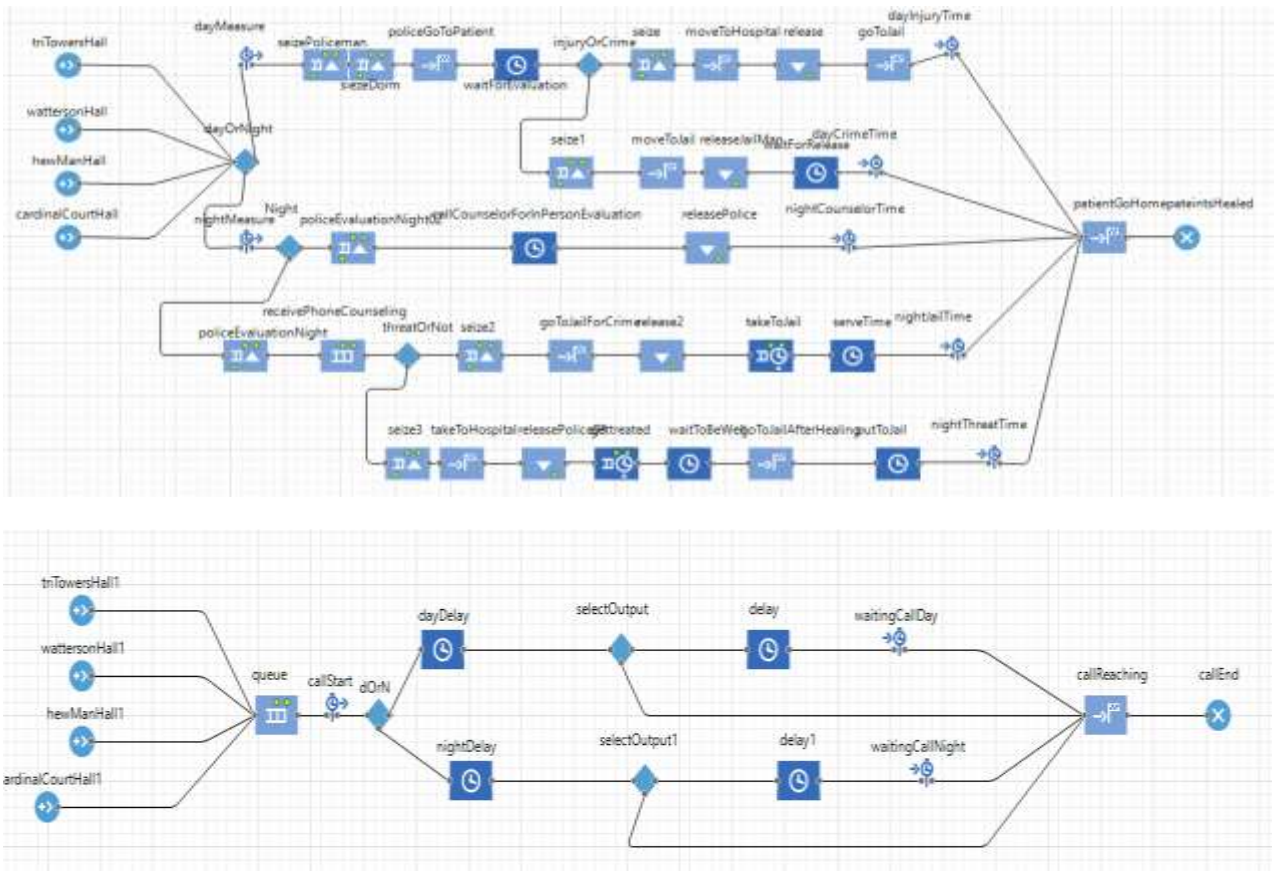


Figure 3. Pre-intervention model logic

Post-Intervention Model

This simulation model in Figure 4 is similar to the first but adds the resource of counselor/psychologist with the officer in the form of the Crisis Intervention Team. The process is more streamlined since treatment can happen immediately upon contact with the person in crisis. It also takes the pressure off the officer in having to make some of these complex decisions, reducing the potential for use of force while on the job. This simulation also has an additional resource of a mental health in-patient center for individuals to utilize when they need additional resources that the hospital/jail cannot provide. This option would be considered the higher standard for the patient over the hospital or jail because they can receive long-term care at the facility.

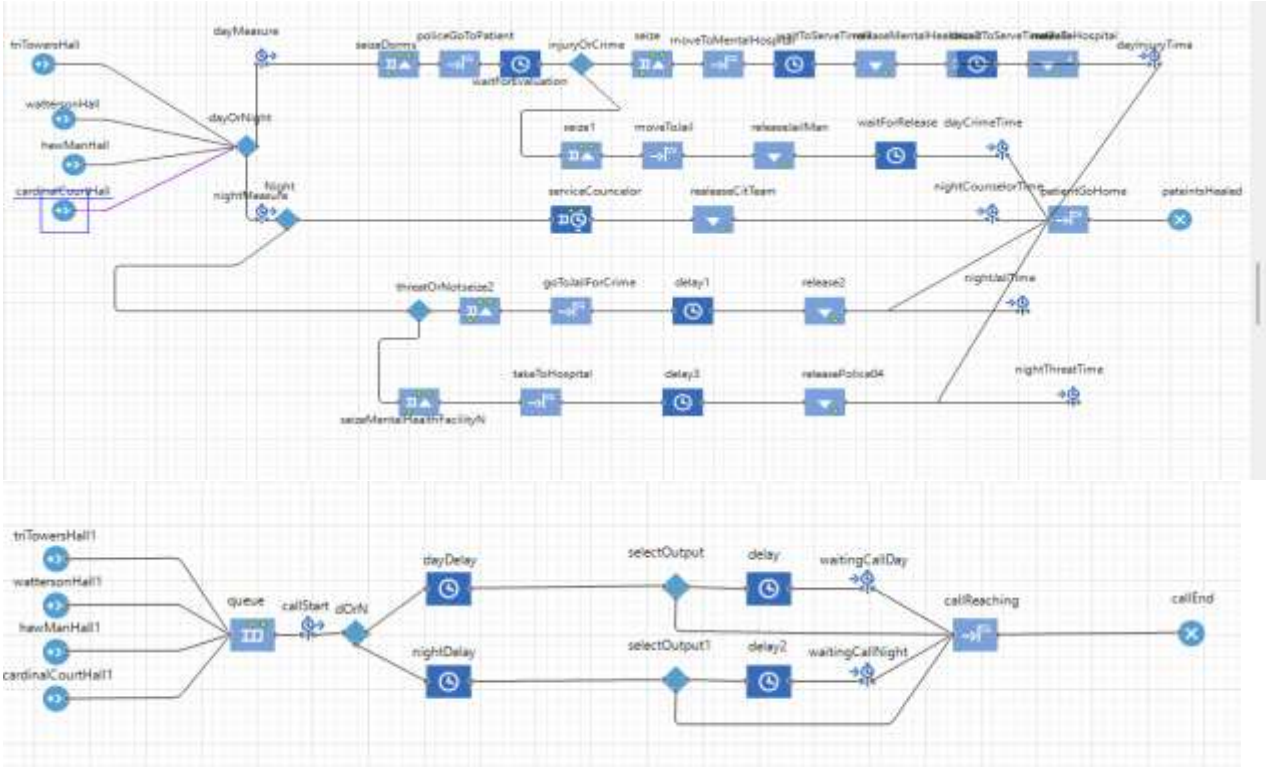


Figure 4. Post-intervention model logic

OUTCOMES

To obtain data, both simulations ran in Any Logic for over 9000 hours, a little over 1 year, to get measurable data for comparison. The results were calculated using three different metrics. The first metric is the average amount of time for the patient to receive help and find a resolution (whether that is going to the hospital, going to jail, or getting released). This was done with a time measure placed at the time that police are dispatched, and the time measure ends just before the patient is released in the process. On the time plots in Figures 5 & 6 it shows that the average wait times are excessive in the pre-intervention model due to the need to wait to speak with a counselor or psychologist. These averages demonstrate that a patient could wait 1 hour to two and a half to speak with someone or be treated, which is obviously not ideal.

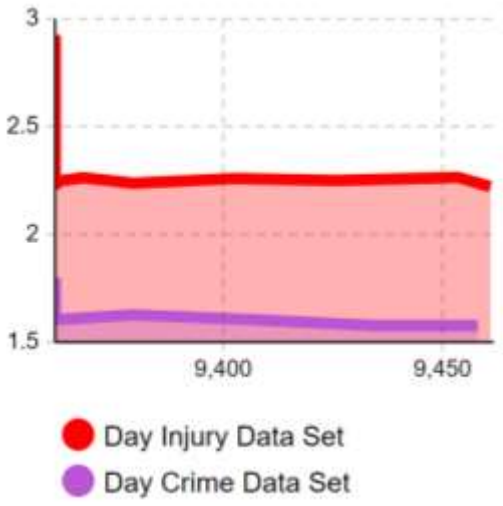


Figure 5 : pre-intervention average time totals

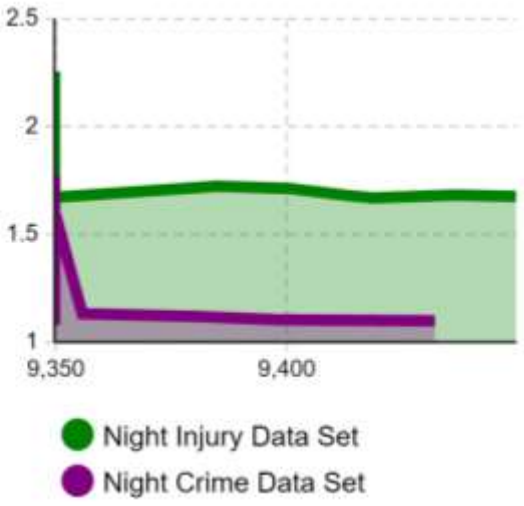


Figure 6: pre-intervention average time totals

In the post-intervention model, the average call times are shown in Figures 7 & 8. The average call times are reduced by at least one hour on average and in some instances even more, meaning a reduction in 40-60% of the overall time spent with officers. This data implies that the patient is getting released sooner, they are being treated in a timely manner, reduced stress on both the officer and patient, and more time for police to spend responding to additional calls.

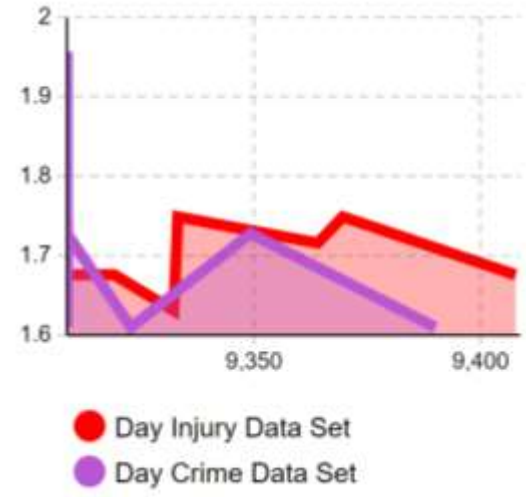


Figure 7: post-intervention average time totals

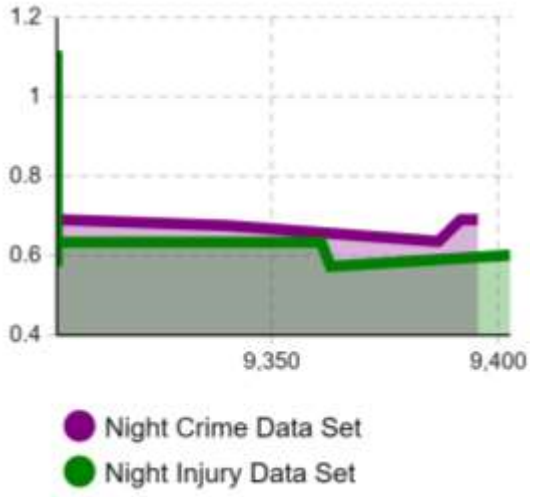


Figure 8: post-intervention average time totals

The second metric used to analyze the process was the utilization of the output resources (jail, hospital, mental health facility). This analysis shows how many people are unnecessarily placed in jail in the pre-intervention model (see Figure 9) and how the process has improved dramatically with the introduction of the CIT (Crisis Intervention Team) unit (see Figure 10). This gives improved outcomes, since patients can get treatment immediately when the CIT unit responds, and the unit can admit patients into a mental health facility and reduce the number of people going to the jail and hospital.

Hospitals and jails only have a limited amount of space, and even as the Midwest Region has seen with COVID, beds in a hospital can easily be in short supply. Figure 10 shows reduction of almost 50 % for both the hospital and jail utilization, which is beneficial for the patient and the community.

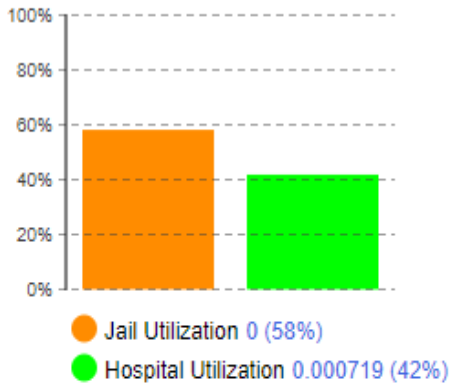


Figure 9 : Pre-Intervention resource utilization

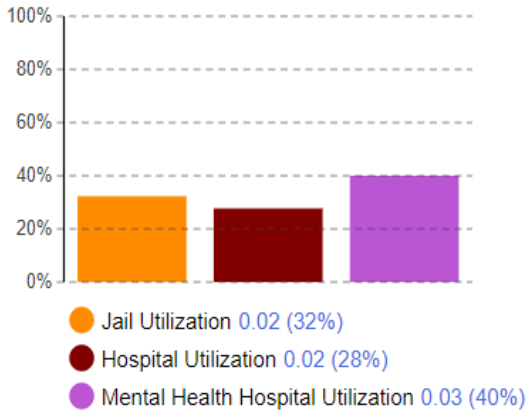


Figure 10: Post-Intervention resource utilization

The last metric was an analysis for the total number of people (patients) who had called for police, received care, and exited the scenario. As shown in the previous metrics, the amount of overall time for patients spent in the post-intervention model was reduced significantly. In this metric, it shows the pre-intervention totals (Figure 11) versus the post-intervention totals (Figure 12), and it shows an overall increase in the total amount of patients treated. This also points to an overall reduction in call times since more patient agents were able to run through the process. The implication of this is not that officers will suddenly have more calls to respond to, but that they will have nearly a third more time available in the post-intervention process to respond to other calls if needed.

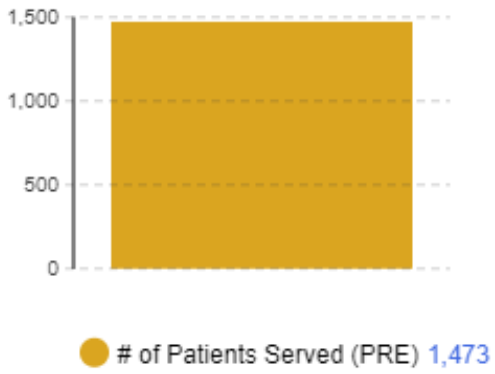


Figure 11: Pre-Intervention Patient totals

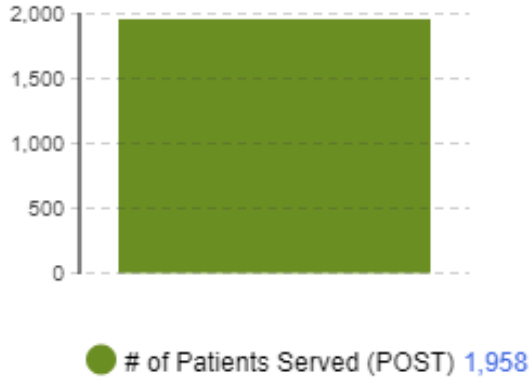


Figure 12: Post-Intervention Patient totals

IMPLEMENTATION AND CONCLUSION

Crisis Intervention is not a new concept and officers have required training on the topic, however it is well known that mental health is of growing concern on a national and local level. Recent events have pointed to the weaknesses in the criminal justice system and suggested that police do not need to be the only first responder in a call for mental health. The idea of implementing a Crisis Intervention Team, in which an in-house counselor or psychologist responds to calls with the officer can present its own challenges. There can be many difficulties in implementing this, especially for some smaller departments, due to funding, lack of resources, buy-in and support, and of course the resources used (time/effort/training) during implementation. The department would need to educate the officers and the community about the benefits of such a program and show how it not only saves time and money for the department but also in the community. Jails and hospitals will use less space on those who simply need mental health treatment, as some of those individuals can be admitted into in-patient mental health facilities. The funding for such a project can come with help from the city, but it is cost-effective when you consider that for the cost of a supervisor's salary, you could put two counselors on the payroll. According to *Glassdoor* the average salary of an entry level counselor is $33K to $51K with an average around $40K (Salary, 2022, n.d.). This could also be the equivalent to one entry level officer's salary, depending on the geographical area. As shown by these analytics, the CIT unit will reduce the time spent on these calls, meaning departments would not feel as overwhelmed or need to hire more officers to deal with these issues. Over time, the benefits of having the counselor would save time spent on calls and with these types of calls increasing, it reduces the need to hire more officers. In fact, money that would have been spent on more officers could then be spent on an additional counselor depending on the department's needs. An additional advantage to the police department is that these counselors not only serve the community but would serve the officers after a traumatic incident. Treating mental health in officers is equally as important and when they frequently put their lives on the line to protect others and are exposed traumatic incidents, they are more likely to suffer from post-traumatic stress disorder, depression, and anxiety, and they are less likely to seek care for these issues (Walden University, 2022).

Some of these factors are difficult to measure and have yet to be measured in a real-world implementation but the measurables in the simulation indicate that the process flow improves

greatly from Pre-Intervention to the Post-Intervention Model. For every department, needs, resources, and funding may vary greatly but the one thing that is clear, is mental health issues are increasing, and police departments will need to find additional resources to deal with it. Even here at the University in the Midwest region, in the past ten years mental health transports have more than doubled. Even if a police department lacks local resources, they can sign a memorandum of understanding (M.O.U.) with a neighboring department to pool resources so everyone can benefit. It only makes sense that a department would benefit from implementing a CIT Unit. Despite the money and resources saved, the bigger point is that it could save lives. Future research on the topic is still needed, but as this idea grows in popularity the community should see the outcomes and benefits. Ideally this model lowers use of force incidents, provides streamlined treatment, reduces the number of people in jail, lessens the amount of people in hospital beds, utilizes local in-patient resources, provides patients with better long-term care, and reduces stress on the patient and officer

REFERENCES

Abramson, A. (2021, July 1). *Building Mental Health into emergency responses*. Monitor on Psychology. Retrieved October 5, 2022, from https://www.apa.org/monitor/2021/07/emergency-responses

Bureau of Justice. (2021, October). *Suicide in Local Jails and State and Federal Prisons, 2000–2019 – Statistical Tables*. Bureau of Justice Statistics. https://bjs.ojp.gov/library/publications/suicide-local-jails-and-state-and-federal-prisons-2000-2019-statistical-tables

Carroll, H. (2016, September). *Serious Mental Illness Prevalence in Jails and Prisons*. Treatment Advocacy Center. Retrieved October 24, 2022, from https://www.treatmentadvocacycenter.org/evidence-and-research/learn-more-about/3695

Herman, C. (2021, June 15). *U Of I Police Department Hiring Social Workers To Handle Mental Health Crisis Calls*. Illinois Newsroom. https://illinoisnewsroom.org/u-of-i-police-department-to-hire-social-workers-to-handle-mental-health-crisis-calls/

ILCS. (2021). *50 ILCS 705/10.17*. Illinois General Assembly. https://www.ilga.gov/legislation/ilcs/documents/005007050K10.17.htm

Rogers, M., McNiel, D., & Binder, R. (2019, September). Effectiveness of Police Crisis Intervention Training Programs. *The Journal of American Academy of Psychiatry and the* Law, Article 1093–6793. https://escholarship.org/uc/item/3670c01r

Wang, L. (2021, June 23). Rise in jail deaths is especially troubling as jail populations become more rural and more female. *Prison Policy Initiative*. https://www.prisonpolicy.org/blog/2021/06/23/jail_mortality/

Reinberg, S. (2022, September 19). Depression Affects Almost 1 in 10 Americans**.** *U.S. News*. https://www.usnews.com/news/health-news/articles/2022-09-19/depression-affects-almost-1-in-10-americans

Salary: Entry Level Counselor (December, 2022). (n.d.). Glassdoor. https://www.glassdoor.com/Salaries/entry-level-counselor-salary-SRCH_KO0,21.htm

Suicide. (2022, June). National Institute of Mental Health (NIMH). https://www.nimh.nih.gov/health/statistics/suicide

Walden University. (2022, November 17). *5 Reasons the Mental Health of Police Officers Needs to Be a Priority*. https://www.waldenu.edu/programs/criminal-justice/resource/five-reasons-the-mental-health-of-police-officers-needs-to-be-a-priority

Power of Community: Engaging Community to Mitigate Terrorism Risk

Abstract

Terrorism poses a major challenge to the government whenever it surfaces in any nation. It degrades the economic, infrastructural and trading systems. Furthermore, it reduces international revenues from tourism and foreign investments. Most importantly, it causes socio-ethnic tension and safety concerns to the nationals. Game-theoretic models have been widely applied to study the strategic interactions between the defender and attacker. In this thesis, we build two subgames between the government and terrorist, and between the government and community, respectively. The community could support either the government or the terrorists. We identify conditions where the community equilibrium strategy is to support the government. Simultaneous games are studied.

Keywords: Business and Terrorism, Game Theory, Community Engagement, First-Mover Advantage

Introduction and Background

In 2002, the Boko Haram group was founded by Mohammed Yusuf in Maidugri, Borno state under the government of Governor Mala Kachalla. Since then, Nigeria has been plagued by intermittent terrorist attacks executed by the Islamist group. “Boko Haram” translates from Hausa, an indigenous Nigerian language, to English as “Western Education is forbidden” or “Book is a sin”. In a self-acclaimed statement, the group also refers to themselves as “Jama'atu Ahlis Sunna Lidda'awati Wal-Jihad”, meaning “People Committed to the Propagation of the Prophet's Teachings and Jihad” (CNN, 2019). The Boko Haram group allegedly withdraws from participation in Nigerian politics as it is against the central Islam notion that the group is structured and regulated by the Sharia (Islamic law). However, it is difficult to categorize their attacks as non-political since they mostly conduct political attacks and attract the attention of the ruling government.

Under the rule of Yusuf, the group was motivated by the “liberation” of the people in the Northern States from Western education, constitutions and institutions. However, when the government intervened and killed Yusuf in 2009, a new leader, Abubakar Shekau, took his place and led the violent uprising of Boko Haram terrorism in its current form. It can be argued that a contributing factor to the increasing violence from the group came from the inauguration of President Goodluck Ebele Jonathan. Other factors could have been negligence from President Jonathan’s administration regarding the importance allocated to counter terrorism measures and elevated poverty and illiteracy levels in the Northern states, which leads to lower standard of living. From observing the strategies used and the frequency of attacks by the Boko Haram terrorist group, it can be gathered that they are a credible threat to concerned nations and should be treated as such. Based on the definition of a credible threat by Dixit and Nalebutt (1991), credible threats contain two components: “a plan of sequenced actions and, the commitment to make the threat credible” (Shan and Zhuang, 2014).

This paper will utilize a game theoretic approach to study and analyze the decision making for three players: the government, community and terrorist with the aim of finding ways to stop the community from supporting terrorists. To the best of the authors’ knowledge, the community has not been studied using game theory for counter-terrorism. The main contributions include examination of the community as a player in counter-terrorism efforts and developing and solving a three-player game involving the government, community and terrorist.

Problem Statement

Potential terrorists might come from the community possibly inspired by the ideology of the terrorist group. The community might play a critical role in counter-terrorism and other government-led initiatives. For instance, the community could make sure its members obeying the safety curfew set by the government for their safety during a pandemic or other emergency situations. Additionally, the community may serve as local aides to the government reporting suspicious incidents though security helplines and centers. This paper will attempt to answer the following questions:

- Under what conditions could the government prevent the community from supporting terrorists?
- How willing is the community to support the government as certain external and internal factors change?

This paper will identify model parameters which might be important for leading to equilibrium involving the decision of the community supporting governmental efforts in counter-terrorism. Next, this study will highlight the feasibility of stopping the community from supporting terrorists. It will be significant in

- Highlighting the importance of the role of community in the mitigation of terrorism, and
- Encouraging the government to develop strategies to better encourage members of the community to support the government rather than the terrorists for the greater good of the nation.

Literature Review

The term "terrorism" has an array of definitions. According to Walter Laqueur, the sole properties of terrorism that are commonly accepted are that terrorism involves both violence and the threat of violence (Laqueur, 1987). The United States Department of State has defined terrorism as "politically motivated violence perpetrated against noncombatant targets by subnational groups or clandestine agents, usually intended to influence an audience" (Charles, 2002). From the definition, three criteria are used to portray the essence of the word. They are political motivation, violence against noncombatants and subnational groups or clandestine agents. Terrorism is considered "politically motivated" since the primary aim of terrorist attacks is to influence the ruling government's decisions to favor the terrorists' cause. The second criteria of "violence against noncombatants" states that the attacks are targeted towards civilians and/or members of the military, who are not on active duty (noncombatants). The third criteria of "subnational groups or clandestine agents" is controversial. Boko Haram attacks cannot fully be described as subnational since their attacks are both within North-Eastern Nigeria, and neighboring countries such as the Republic of Chad, Cameroun and Niger. Also, they can no longer be categorized collectively as clandestine agents except in terms of individual agents since some of their operatives are of known identity. According to the 2019 Global Terrorism Index, Nigeria ranks 3rd with a score of 8,597 indicating a very high terrorism impact on the country (Institute for Economics & Peace, 2019). Numerous perspectives arise when comparisons are made between terrorists and freedom fighters. In this context, demonstrations by the Nigerian Niger Delta Militants are compared to those by the Boko Haram terrorist group. Both are being labelled as terrorists by some schools of thought. Although it can be gathered that both groups are motivated by the pursuit of liberation, and achieve their goals through violence and forceful moves, there is a major difference between both groups.

Freedom fighters are motivated to act against the ruling government when the government deprives the people of fundamental necessities that are standards for humane living conditions. On the other hand, terrorism acts are calculated acts of destruction on the people and infrastructure conducted with the sole purpose of causing deliberate harm to both the nation and the nation's population (Ochoche, 2013).

Furthermore, organized crimes and terrorism share notable differences. Although both organized crimes and terrorism share some similarities such as the need for money (to launch attacks), some differences exist between them. These differences include the aim of organized crimes and terrorism being to yield economical profits and affect political activities, respectively (Passas, 2005). While organized crimes mostly pursue anonymity, terrorism seeks media coverage of attacks, where the concerned terrorist group takes the responsibility for an attack as a form of intimidation and as a method to attract new followers. Another distinguishing trait between organized crimes and terrorism is that terrorist attacks are ocassionally planned to target prominent governement agencies or nationally recoginsed commemoration dates and anniversaries. The potential threat that terrorism poses to the targeted community worsens with time. According to the United Nations High Commission for Refugees (UNHCR), over 3.3 million people have been displaced since the Boko Haram terrorists' attacks heightened in 2014. Within the 3.3 million, over 2.5 million of the internally displaced persons (IDPs) are reportedly in North-Eastern Nigeria, whereas over 550,000 are reportedly in the Republic of Chad, Cameroon and Niger. Also, as of 2014, there were approximately 240,000 refugees in Nigeria, Cameroon, Chad and Niger altogether (United Nations High Commissioner for Refugees, 2001-2020).

Generally, the community plays a vital role in the mitigation and prevention of crimes and pandemics. This could be achieved through the formation of a Neighborhood Watch group. A Neighborhood Watch is a police-enabled program that encourages members of the community to help protect themselves and support the police. According to John Woodhead, the Vice President and Executive Chairman of Queensland Neighborhood Watch Association Inc., a major issue with all association branches was receiving cooperation from the local police departments since the degree of support varied in all states. It was reported that the establishment of the Neighborhood Watch program reduced the rate of burglaries and crimes in general in the Queensland, Australia, and in the other states (Woodhead, 1990). Also, members of the community formed close-knit relationships with the police as they worked together towards achieving common goals. However, recently the Neigbourhood Watch seems not as effective. According to Jenny Fleming (2005), the Neigbourhood Watch despite having large numbers of volunteers are only slightly reducing crimes in areas with low crime rates. However, it could be due to the fact that the studied areas already have low-crime rates. Further examining the role of the community in the fight against terrorism in Nigeria leads us to the events that occurred on May 13, 2014. Boko Haram terrorists, who attacked three villages in Maiduguri, Borno state, were met with resistance from the community leading to the deaths of over 200 Boko Haram terrorists (CNN, 2019).

The important role of the community in detering adverse activities in their nation cannot be overstated. While highlighting the importance of the community's role in the prevention of global issues, we examine community responses and roles in deterring a recent global pandemic, the spread of the COVID-19 disease, social distancing, increased personal and household hygiene, observing self-quarantine, obeying the federal quarantine orders, and obeying the orders put in place by essential facilities are some of the ways that the community helped reduce the coronavirus spread (Harvard Health Publishing, 2020). Another highly important role that members of the

community play is in supporting one another in crisis. Having support is fundamental for members of affected communities especially for children and young adults, who may not fully comprehend an on-going crisis. Community supports in crisis can be discounted supplies in neighborhood stores, aiding neighbors whenever needed, and keeping one another informed on the new changes or policies (Boudreau, 2020).

The government mainly focuses on deterrence as a form of terrorism mitigation. They deter the attacks by instilling doubts and fears within the nationals regarding terrorism and the punishments for acts of terrorism if prosecuted. However, deterrence has not always been effective. In fact, it discourages innocent members of the community from providing tips as they fear being punished as severely as the terrorists if they were assumed to be members of the terrorist groups. Richardson et al. (2007) offered three strategies to terrorism mitigation that were considered "superior strategies" to deterrence. The proposed strategies influence the terrorism by "decreasing the utility of terrorism to terrorists" and by "attempting to increase the opportunity cost." The strategies are polycentricity, diffusing the attention of the media and offering positive incentives. Polycentricity focuses on creating more than one core/center for the system. In a case where one part of the system is attacked or negatively impacted, the other parts can take control. This is expected to reduce the vulnerability of the overall system and thus decrease the utility of terrorism to terrorists. The concept of diffusing the media suggested by the authors is by the government curtailing credit given to terrorist groups for acts of terrorism in the media. This is expected to reduce the utility of terrorism to the terrorists since gaining public recognition is a major goal for terrorists. However, this strategy is a double-edged sword since it creates media confusion in the community. Individuals will not be completely informed on whether there are more than one existing terrorist groups on attacks or the extent to which their safety is threatened, which can lead to further distrust for the media and government. The last strategy involves giving positive incentives to the terrorists as a mitigation method by "reintegrating terrorists and providing access to the political process, and welcoming repentants" (BBC News, 2020). Although this strategy might be welcoming to the terrorists since it includes access to political processes (which is usually the aim of the attacks) and pardon to "repentants", a number of questions arise and must be considered. How does the government decide on the extent of political access to provide to the terrorists? How does the government ensure that "repentants" truly have repented? Do the "repentants" face the judicial system? Will the community accept "repentants" back after the damage they have already caused? These questions are extremely critical to the success of this strategy to increase the opportunity cost.

To provide a closer look at the driving forces behind the Boko Haram Insurgency, Alozieuwa (2012) examined various theories, including the Relational/Vengeance Theory, Human Needs/Socio-Economic Perspective, Political Feud Perspective, Islamic Theocratic State Theory and the Conspiracy Theory. After careful examination, it was concluded that originally the Boko Haram attacks were solely driven by religions but evolved to being influenced heavily by politics. In an attempt to develop thorough understanding of the political implications of Boko Haram attacks on Nigeria, Chukwurah et al. (2015) conducted a study. It was gathered that there was an increment of Boko Haram activities as a president from Southeastern Nigerian was inaugurated in 2010. Six sectors were examined during this research. They are the tourism, transportation, infrastructural, commercial, core service and agricultural sectors. The transportation agencies experienced a drastic decline in passengers travelling to the North Nigeria with the uprising of Boko Haram attacks. Tourism from the North, which produced approximately 80 billion naira

yearly (the highest in Nigeria), experienced a standstill and then a decline. Infrastructure have been severely damaged, and both foreign and local contractors assigned to be engaged in the rehabilitation of these infrastructure have fled those states due to concerns for their safety. The commercial sector, which comprised of corner shops, indigenous markets and small-to-medium businesses, experienced approximately a 73.7% decline and people live on their savings or migrate. The core services such as schools, hospitals, hotels and parks are moving their businesses to other parts of the country. The agricultural sectors are affected since farm lands are no longer being tended to and farmers are concerned for their safety.

Access to improved education, employment, infrastructures such as healthcare facilities and thus a better standard of living might mitigate terrorism in the Northern parts of Nigeria since the poverty and illiteracy levels are relatively high in those regions. However, the members of this terrorist group are in fact exposed to Islamic education, which when misinterpreted could influence them to develop a sense of misguided purpose to facilitate the forceful spread of the religion by any necessary means. Adesoji (2010) found that exposure to Western education was not the solution since "For them, it is a passion, a belief system. I do not believe that exposure to formal western education is the key to mitigate these terrorists". On the other hand, Adetoro (2012) suggested that thorough state reforms focusing on infrastructural, structural and considerable poverty relief programs would curtail Boko Haram attacks and other ethno-religious crises in Nigeria. This was concluded after a study was conducted considering poverty and political alienation as major indicators for the Boko Haram Insurgency. Although poverty might be a major factor to consider, the Boko Haram terrorist group might be solely driven by greed and personal vendetta. These are not strange grounds in Nigeria especially when the pursuit of power and political motives are involved.

Mitigating acts of terrorism could be facilitated by considering all factors available and necessary. The game-theoretic approach is a mathematical method of analysis used to understand the interactions between two or more players. This approach considers possible actions from each player and then uses payoffs to determine the optimality of the move. Generally, the government should combat terrorists as that serves the society far more than when no action is taken; however, this is not always the case (Daniel, 2005). Ochoche (2013) performed game theoretic studies using four models: 1, 2A, 2B and 3. In the first model, which is zero-sum, the equilibrium strategy for both the terrorist group and the government were to terrorize and combat, respectively. The second model considered a scenario where the home zone states increased the cost of terrorism and initiated cease fire incentives. The equilibrium strategy remains the same. Model 2B considered increasing the cost of terrorism (a decrease in the payoff). This resulted in cease fire and combat being the dominant strategies for the terrorist group and government, respectively. Ochoche's fourth model analyzed the condition that the terrorist group has an understanding with the home zone state and moves operations to other states excluding that home zone state from attacks. The dominant strategies are to terrorize and compromise for the terrorist group and state government, respectively. This approach endangers another state and does not solve the terrorism problem for the nation or the home zone state. Also, there are no guarantees that the terrorists will continue to comply with the agreement. However, it is unclear how an agreement is reached. The proposed counter-terrorism strategies include increasing the cost of terrorism, initiating cease fire, and hindering alliance between state government and terrorist groups. This paper only considers simultaneous games between the government and the terrorist and the payoffs are determined with stylized values.

Game Theory Model Formulation

Game theory is a mathematical method that studies the strategically interdependent behavior between players (Barron, 2013). Interdependence refers to the fact that an action from one player affects the other and vice versa. For this paper, three players are considered: the government, community and the terrorists. The payoffs will be determined based on literature review and analysis of their interests.

The government has two options of combating and engaging the community. The terrorist group has the options of attacking and not attacking. The community takes the role of supporting either the government or the terrorist group. The game tree is presented in Figure 1.

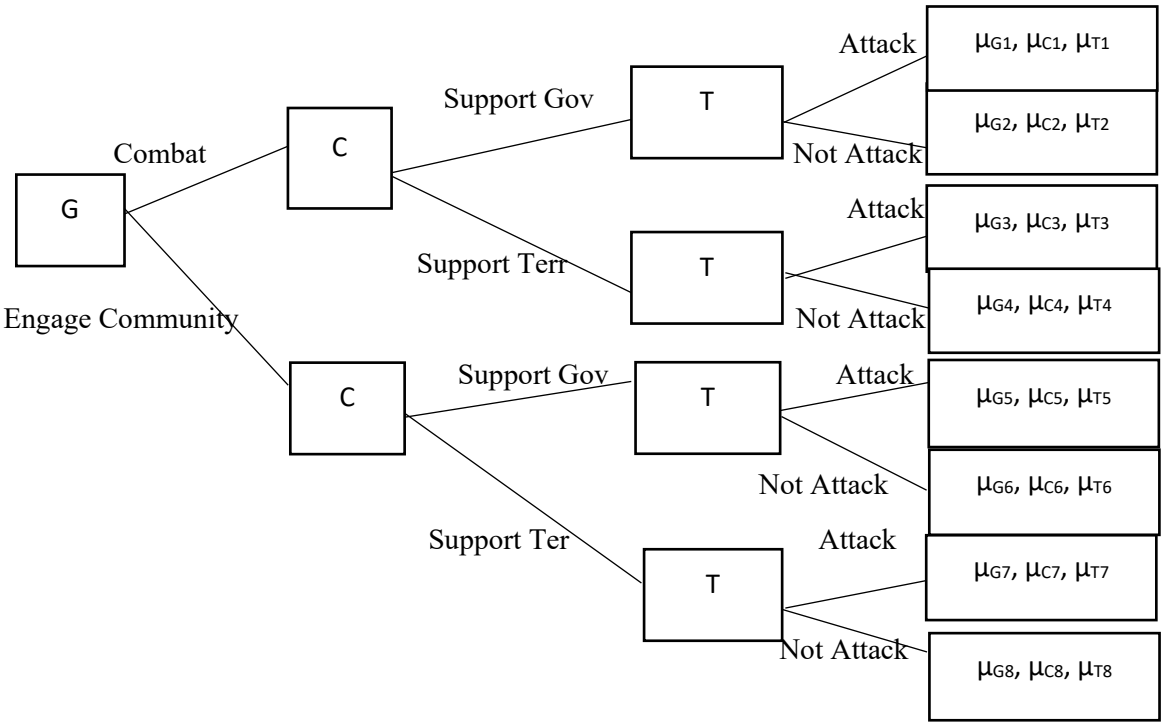


Figure 1. Game tree of the three players: the government, terrorist, and community with their payoffs.

Notation

The n-tuple (player) game is mathematically presented as $G = (N, S, \mathrm{m})$, where $N = \{1,2,3,\ldots,n\}$ is the set of players, $S = \{S_1, S_2, S_3 \ldots, S_n\}$ and $\mathrm{m} = \{\mathrm{m}_1, \mathrm{m}_2, \mathrm{m}_3, \ldots, \mathrm{m}_n\}$ are the strategy and payoff sets, respectively.

Every player can perform an action of A_i

Assume $m_i = (a_i^0, a_i^1, a_i^2, \ldots, a_i^{n-1})$ *to be the sequential actions taken by player* $i \in N$.

The zero-sum game in a tactical game is one that for all $s = \{s_1, s_2, s_3, \ldots, s_n\} \in S$,

$$\mathrm{m}_1(s) + \mathrm{m}_2(s) + \mathrm{m}_3(s) + \ldots + \mathrm{m}_n(s) = 0$$

All players play to maximize their payoffs. Two strategies s_i and s_i^0 are given so that in any strategy combination, the result from s_i is greater than that from s_i^0. In a set of $s_1, s_2, s_3, \ldots, s_n$, if

$$\mu_i(s_1, s_2, s_3, s_i \ldots, s_n) \geq \mu_i(s_1, s_2, s_3, s_i^0 \ldots, s_n)$$

$s_i \in S$, where s_i represents player i's strategies

Nash equilibrium is reached at $\boldsymbol{s} = \{s_1, s_2, s_3, \ldots, s_n\}$

Justification of the Government's Choices

Government can choose to either heighten security or engage the community. If the government chooses to heighten security, then they must consider defense strategy against the terrorists, cost of the defense, and the impact factor. Generally, the government aspires to target the terrorists while minimizing damages to the community including infrastructure damages, casualty, societal tie disruptions, and so on.

According to the United Nations, the government heightens security by securing both domestic and international borders, constricting financial regulations, increasing the involvement of police authorities, improving the criminal justice system, and establishing legal alliance with other countries with terrorist threats to help convict terrorists in their courts (Smith, 2020). By defending, the government may monitor websites and online contents, to take down hate speech and suspicious activities. The Dutch government has implemented numerous actions to strengthen its defense against terrorism. The Royal Netherlands Air Force surveys the Dutch airspace on a 24 hours basis. Also, the police closely monitor people, who they suspect to be terrorist threats to the society (Government of the Netherlands, 2016-2020). The "Counterterrorism Alert System" established by the Dutch government notifies the government and primary sectors about terrorist threats (2016-2020). These sectors including drinking water, and energy companies.

Other impactful strategies by the government to form defense against terrorism can be exemplified through the North Atlantic Treaty Organization (NATO). NATO narrows efforts towards terrorism mitigation through improving awareness of terrorist threats, setting up capabilities to prepare and adequately respond to threats, and forming alliances/partnerships with neighboring countries and other international actors (NATO, 2019).

According to Zycher (2003), annual costs of deaths and injuries from moderate, severe, and nuclear cases in the United States were estimated at $11 billion, $183 billion, and $465 billion, respectively, before the September 11 attacks. These values are based on the estimates that one life is worth $4 million, and one injured individual is worth $40,000. However, after the September 11 attacks, government spending for reconstructing, humanitarian activities, defense, economic aid, and domestic security functions have been estimated at $95 billion.

The government can engage the community by establishing community enrichment programs, and protecting and rewarding informants. The Office for Victims of Crime (OVC) in the U.S. Department of Justice provides 24/7 crisis counselling services in English and Spanish to victims of terrorism and mass violence through the "Terrorism and Special Jurisdictions Program" (Office for Victims of Crime, 2020). Also, OVC provides victim compensation to the affected states for the welfare of the affected within the state. Other organizations that provide support are "The Dougy Center", "National Organization of Parents of Murdered Children", and "VictimConnect" (Office for Victims of Crime, 2020). The United States' government through the Federal Bureau of Investigation (FBI) established the FBI Counterterrorism Center in 1996 (Watson, 2002) for combating terrorist activities on both domestic and international levels.

The Nigerian government has made efforts to engage the community through providing incentives to the community for disengaging in violent acts. Recently, the Zamfara State Governor, Bello Matawalle, has offered two cows to the indigenous people for every AK-47 or weapon returned to the government (BBC News, 2020). Zamfara is a state in northwestern Nigeria with approximately 67.5% people living in poverty (BBC News, 2020). A lucrative business in Zamfara State is farming, especially animal herding, and cows are highly valued by the average herdsman. Also, over 8,000 people have been killed through crimes related to terrorism, theft, and inter-ethnic tension in Zamfara and its neighboring states, which makes this incentive not only relevant but also necessary. Another way that the government can engage the community is through organizations such as the U.S. Intelligence Community (IC) Civilian Joint Duty Program, which offers civilians with professional opportunities to strengthen collaboration with the government and community. Also, the program provides the participating civilians with enhanced career prospects through trainings and exposure to the processes involved in intelligence (Office of the Director of National Intelligence, 2020).

Exploring Different Options for the Community

The community can choose either to support the government or support the terrorist. The community can support the government by reporting suspicious activity to the police authorities, become involved in campaigns organized by the government to influence the children and youth, and refrain from voluntary recruitment by the terrorist. In 2007, the U.S. government established the Nationwide Suspicious Activity Reporting (SAR) Initiative (NSI), which is a collective effort by the FBI, Department of Homeland Security, state, local, tribal, and territorial law enforcement partners (Joint Counterterrorism Assessment Team, 2020). This initiative informs on how to safely report suspicious activities to the appropriate authorities.

Payoff Functions

To derive a better understanding of the players, their decisions, and outcomes and thus payoff functions, their definitions and other notations are outlined in Table 1.

Table 1. Notation for players, decisions, payoff, and their descriptions.

Player	*Description*
G	Government
C	Community
T	Terrorist
Decision Variable	*Description*
{H, EC}	Government heightens security or engages community
{SG, ST}	Community supports government or terrorist
{A, NA}	Terrorist attacks or not
Government's Payoff	
Parameter	*Description*
L_G	Impact factor to government: cost of damages to the government after defense
I_G	Impact to government for engaging the community when the terrorist attacks

J_G	Loss to government when government heightens security and community supports terrorist
H_G	Loss to government when government engages community and community supports terrorist
V_G	Benefit to government when government engages community and community supports government
C_D	Cost of defense to the government
C_E	Cost of engaging the community to the government
Community's Payoffs	
Parameter	*Description*
S_E	Impact to community when government engages community and community supports government
S_W	Impact to community when government engages community and community supports terrorist
S_G	Impact to community for supporting government when government heightens security
S_T	Impact to community for supporting terrorist when government heightens security
I_C	Impact to community for supporting government when terrorist attacks
T_C	Impact to community for supporting terrorist when terrorist attacks
X_C	Benefit to community for supporting government when terrorist does not attack
P_C	Loss to community for supporting terrorist when terrorist does not attack
Terrorist's Payoffs	
Parameter	*Description*
L_T	Impact factor to terrorist: cost of damages caused by terrorist after attack
O_T	Impact to terrorist when government engages community and terrorist attacks
K_T	Impact on terrorist when community supports government and terrorist attacks
B_T	Benefit to terrorist when community supports terrorist and terrorist attacks
MC	Media coverage for terrorist attacks
C_A	Cost of attack to the terrorist
Other Notation	
{Y, N}	Possible equilibrium or not a possible equilibrium, respectively
{μ}	Assigned payoff

Considering the Two Player Game between Subgames

We first consider two-player sequential and simultaneous games between the government and terrorist, the government and community, and the community and terrorist, respectively.

Government and Terrorist

Figure 2 shows the sequential game between the government and terrorist. In this sequential game, the government makes the first move, and the terrorist is aware of the government's move before their move is made. When the government heightens security (H), the payoff is $-\mathrm{L_G} - \mathrm{C_D}$ or $-\mathrm{C_D}$ depending on if the terrorist attacks or not, respectively. When the government engages the community (EC), the payoff is $-\mathrm{I_G} - \mathrm{C_E}$ or $-\mathrm{C_E}$ depending on if the terrorist attacks or not, respectively. On the other hand, the terrorist payoff is $L_T + MC - C_A$ or 0 depending on if they attack or not when the government heightens security. Similarly, the terrorist payoff is $L_T + O_T + MC - C_A$ or 0 depending on if they attack or not when the government engages the community.

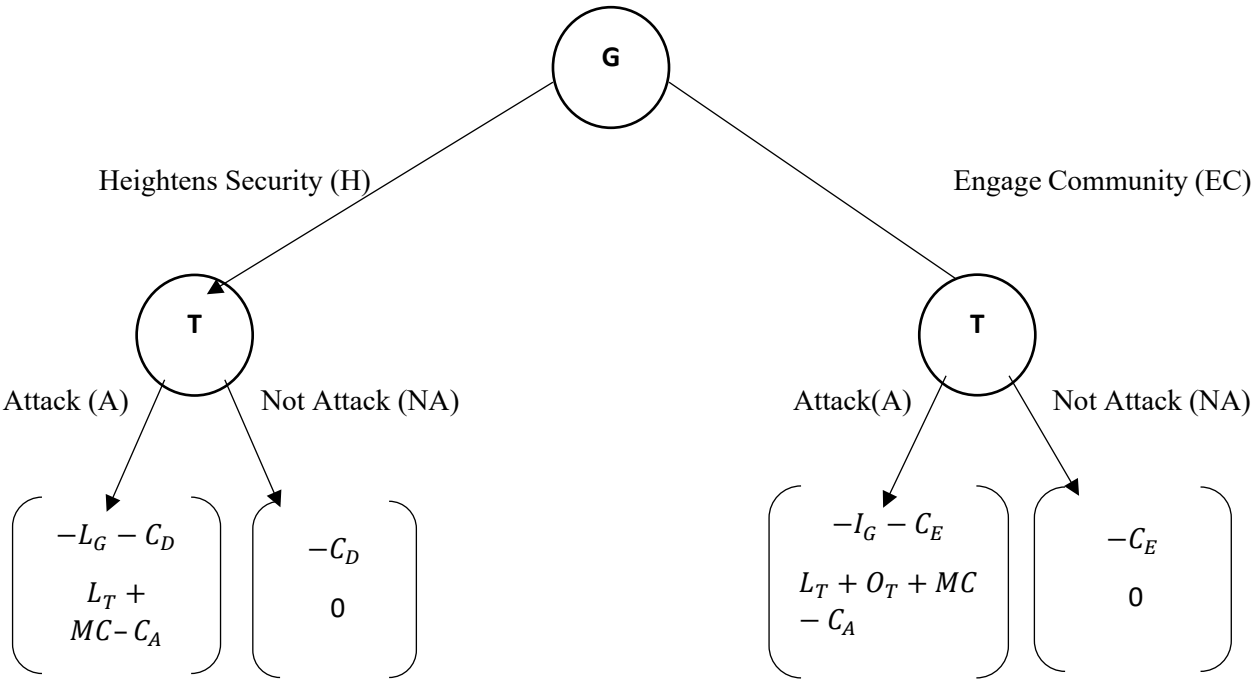


Figure 2. Game tree of the sequential subgame between the government (G) and terrorist (T).

Table 2. Possible equilibria for subgame between government and terrorist

Case	Strategy	Payoff	Condition
1	[H, A]	$[-L_G - C_D, L_T + MC - C_A]$	$\{C_A < L_T + MC, C_A < L_T + O_T + MC, C_E > L_G + C_D - I_G\}$
2	[H, NA]	$[-C_D, 0]$	$\{L_T + O_T + MC < C_A < L_T + MC, C_E > L_G + C_D\}$ *or* $\{L_T + MC < C_A < L_T + O_T + MC, C_E > C_D - I_G\}$ *or* $\{C_A > max\{L_T + MC, L_T + O_T + MC\}, C_E > C_D\}$
3	[EC, A]	$[-I_G - C_E, L_T + O_T + MC - C_A]$	$\{C_A < L_T + MC, C_A < L_T + O_T + MC, C_E < L_G + C_D - I_G\}$

4	[EC, NA]	$[-C_E, 0]$	$\{L_T + O_T + MC < C_A < L_T + MC, C_E < L_G + C_D\}$ or $\{L_T + MC < C_A < L_T + O_T + MC, C_E < IG - C_D\}$ or $\{C_A > max\{L_T + MC, L_T + O_T + MC, C_E < C_D\}$

Solution

In the sequential game, from solving the subgame between the government and terrorist, all four cases present feasible equilibria provided the conditions are met (as shown in Table 2). They are:

1. When the cost of attack is low ($C_A < min\{L_T + MC, L_T + O_T + MC\}$), and the cost of engaging the community to the government is high ($C_E > L_G + C_D - I_G$), the government will heighten security and the terrorist will attack.
2. When the cost of attack is medium ($L_T + O_T + MC < C_A < L_T + MC\ if\ O_T < 0; or\ L_T + O_T + MC < C_A < L_T + MC\ if O_T \geq 0$), the cost of engaging the community to the government is high ($C_E > L_G + C_D$), or when the cost of attack is medium ($L_T + MC < C_A < L_T + O_T + MC\ if O_T \geq 0$) and the cost of engaging the community to the government is relatively high ($C_E > C_D - I_G$), or when the cost of attack is high ($C_A > max\{L_T + MC, L_T + O_T + MC\}$) and the cost of engaging the community to the government is high ($C_E > C_D$), the government will heighten security and the terrorist will not attack.
3. When the cost of attack is low ($C_A < \max\{L_T + MC, L_T + O_T + MC\}$), and the cost of engaging the community to the government is low ($C_E < L_G + C_D - I_G$), the government will engage the community and the terrorist will attack.
4. When the cost of attack is medium ($L_T + O_T + MC < C_A < L_T + MC\ if\ O_T < 0$) and the cost of engaging the community to the government is low ($C_E < L_G + C_D$), or when the cost of attack is medium ($L_T + MC < C_A < L_T + O_T + MC\ if O_T \geq 0$) and the cost of engaging the community to the government is relatively low ($C_E < L_G + C_D$), or when the cost of attack is high ($C_A > max\{L_T + MC, L_T + O_T + MC\}$), and the cost of engaging the community to the government is low ($C_E < C_D$), the government will engage the community and the terrorist will not attack.

However, considering that this subgame is a sequential game where the terrorist is aware of the government's moves before their move is selected, the most desirable outcomes for the government will be one that minimizes the impact of the damage to the community by the terrorist, and defends the community and nation as a whole. Hence, the most desirable equilibrium for the government will be [H, NA] and [EC, NA], where the terrorist does not attack provided the conditions in Table 2 for these cases are met. [EC, A] will not be a desirable equilibrium for the government since when the government engages the community, the terrorist will prefer to attack to obtain a higher payoff than otherwise, which is not desirable to the government under the given conditions. [H, A] will not be a desirable equilibrium for the government since when the government heightens security, the terrorist will prefer to attack to obtain a higher payoff than otherwise, which is not desirable to the government under the given conditions.

For the terrorist, the most desirable equilibrium is the one where the most damage is achieved. Hence, the desired equilibria for the terrorist will be [H, A] and [EC, A], where the terrorist attacks

provided the conditions in Table 2 for these cases are met. [EC, NA] will not be a desirable equilibrium for the terrorist since if the government chooses to engage the community, the terrorist will prefer not to attack to receive a higher payoff than otherwise under the given conditions. [H, NA] will not be a desirable equilibrium for the terrorist since if the government chooses to heighten security, the terrorist will prefer not to attack to receive a higher payoff than otherwise under the given conditions.

Next, considering a simultaneous game where neither player knows the moves of the other player, both players' moves are made at the same time solely based on the best outcome for the individual player regardless of the other player's move as shown in Table 3. To determine the conditions for the equilibrium, we use best response analysis to find the equilibrium. First, we evaluate the player's payoffs from the government's perspective, then we evaluate the payoffs from the terrorist's perspective.

Table 3. Simultaneous subgame between the government (G) and the terrorist (T).

	T	
G	Attack (A)	Not Attack (NA)
Heighten Security (H)	$[-L_G - C_D, L_T + MC - C_A]$	$[-C_D, 0]$
Engage Community (EC)	$[-I_G - C_E,\ L_T + O_T + MC - C_A]$	$[-C_E, 0]$

Table 4. Possible equilibria for the simultaneous subgame between government and terrorist

Cases	**Strategies**	**Payoffs**	**Conditions**
1	[H, A]	$[-L_G - C_D, L_T + MC - C_A]$	$\{C_A < L_T + MC, C_A < L_T + O_T + MC, C_E > L_G + C_D - I_G\}$
2	[H, NA]	$[-C_D, 0]$	$\{C_A > max\{L_T + MC, L_T + O_T + MC\}, C_E > C_D\}$
3	[EC, A]	$[-I_G - C_E, L_T + O_T + MC - C_A]$	$\{C_A < L_T + MC, C_A < L_T + O_T + MC, C_E < L_G + C_D - I_G\}$
4	[EC, NA]	$[-C_E, 0]$	$\{C_A > max\{L_T + MC, L_T + O_T + MC, C_E < C_D\}$

Solution

In the simultaneous game, from solving the subgame between the government and terrorist, all four cases present feasible equilibrium provided the conditions are met as shown in Table 4. They are:

1. When the cost of attack is low ($C_A < \min\{L_T + MC, L_T + O_T + MC\}$), and the cost of engaging the community to the government are high ($C_E > L_G + C_D - I_G$), the government will heighten security and terrorist will attack.

2. When the cost of attack is high ($C_A > \max\{L_T + MC, L_T + O_T + MC\}$) and the cost of engaging the community to the government are high ($C_E > C_D$), the government will heighten security and terrorist will not attack.
3. When the cost of attack is low ($C_A < \min\{L_T + MC\ ,\ L_T + O_T + MC\ \}$) when the community supports either the government or the terrorist, and the cost of engaging the community to the government is low ($C_E < L_G + C_D - I_G$), the government will engage the community and terrorist will attack.
4. When the cost of attack is high ($C_A > max\{L_T + MC, L_T + O_T + MC\}$) and the cost of engaging the community to the government are low ($C_E < C_D$), the government will engage the community and terrorist will not attack.

Considering that this subgame is a simultaneous game, where the terrorist is not aware of the government's move before their move is selected, and both players select their moves at the same time. Similar to the sequential game, the most desirable outcomes for the government will be the one that minimizes the impact of the damage to the community by the terrorist and defends the nation as a whole. Hence, the most desirable equilibrium for the government will be [H, NA] and [EC, NA] where the terrorist does not attack provided the conditions in Table 4 for these cases are met. [EC, A] will not be a desirable equilibrium for the government since when the government engages the community, the terrorist will prefer to attack to get a higher payoff than otherwise which is not desirable to the government under the given conditions. [H, A] will not be a desirable equilibrium for the government since when the government heightens security, the terrorist will prefer to attack to get a higher payoff than otherwise which is not desirable to the government under the given conditions.

For the terrorist, the most desirable equilibrium is the one where the most damage is achieved. Hence, the equilibrium for the terrorist will be [H, A] and [EC, A], where the terrorist attacks provided the conditions in Table 4 for these cases are met. [EC, NA] will not be a desirable equilibrium for the terrorist since if the government chooses to engage the community, the terrorist will prefer not to attack to receive a higher payoff than otherwise under the given conditions. [H, NA] will not be a desirable equilibrium for the terrorist since if the government chooses to heighten security, the terrorist will prefer not to attack to get a higher payoff than otherwise under the given conditions.

Comparing the Possible Equilibria for the Sequential and Simultaneous Subgame between the Government and Terrorist

Comparing the payoffs for both the sequential and simultaneous subgames between the government and terrorist, the conditions for possible equilibrium for cases 1 and 3 (the most desirable outcomes for the terrorists) are the same. However, for cases 2 and 4 (the most desirable outcomes for the government), there are more conditions in the sequential game than in the simulations game. This indicates that the government is more likely to reach their desired outcome in a sequential game than in a simultaneous game. Hence, we recommend that the government announce their strategy to the terrorists rather than keep their strategy unknown to the terrorist to better reach their desired outcome.

In Figure 3a, when $O_T \geq 0$, the cost of attack for the terrorist is low, and the cost of community engagement on the government when the terrorist attacks is high, and might become detrimental to the government if engaging the community is not successful.

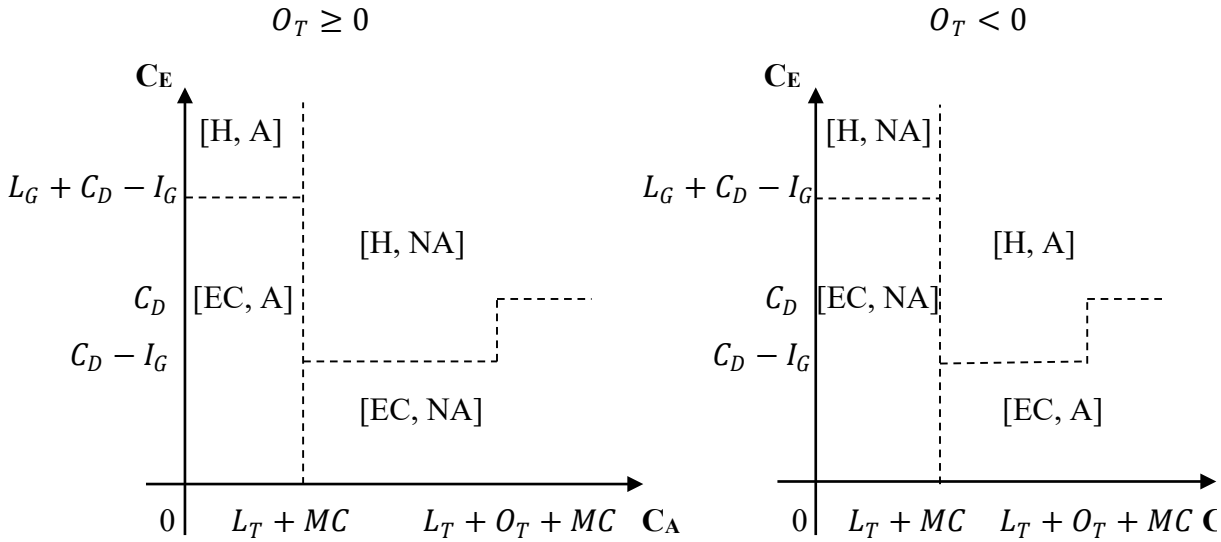


Figure 3. Four possible equilibria (as specified in Table 2) for subgame between government and terrorist as a function of C_A and C_E. The strategies in the brackets are for the government and terrorist respectively.

In Figure 3b, when $O_T < 0$, the cost of attack for the terrorist is high, and the cost of community engagement for the government when the terrorist attacks is relatively low, the government has a higher chance of success with community engagement, while the terrorist will likely not attack when cost of attack is high

Conclusion

In this research, game theory was used as a means to analyze the interactions between the government, terrorists and the community with regards to terrorism mitigation. One subgame between the government and terrorist was developed and solved. Both sequential and simultaneous games are compared. We identified possible equilibria and their leading conditions in both orders of moves. We also analyzed the most desirable equilibria. We compared the sequential and simultaneous games and confirmed the first mover advantage in the sequential game between the government and terrorist (when the government moves first) supporting previous research (Zhuang and Bier 2007). While studying the two-player sequential and simultaneous subgames games, the results show that all three players have more conditions to reach possible equilibrium in the sequential game rather than in the simultaneous game. This leads to the first conclusion that the government, make their strategy publicly available to the community, and terrorist in order to receive their desired outcomes rather than keeping it undisclosed.

This topic is relevant for mitigating terrorism since terrorism is a growing world challenge, which endangers innocent citizens, increases crime rates, and deteriorates the economy of concerned nations. The community being the most vulnerable sub-unit of the nation could experience high displacement into already relatively saturated communities due to terrorism in some countries. The rise of lone wolf terrorist acts suggests the importance of engaging the community so that they are likely to support the terrorist or becoming a terrorist. Therefore, it is important to find efficient ways to stop the community from supporting terrorists out of necessity to survive.

References

Adesoji, A. (2010). The Boko Haram Uprising and Islamic Revivalism in Nigeria. *Africa Spectrum* , 100-104.

Adetoro, R. A. (2012, November-December). Boko Haram insurgency in Nigeria as a symptom of poverty and political alienation. *Journal of Humanities And Social Science, 3*(5), 21-26.

Alozieuwa, S. H. (2012). Contending Theories on Nigeria's Security Challenge in the Era of Boko Haram Insurgency. *Peace & Conflict Review, 7*(1), 1-8.

Barron, E. N. (2013). *Game Theory : An Introduction* (2 ed., Vol. 1). John Wiley & Sons, Incorporated.

BBC News. (2020, July 17). Nigeria's Zamfara state offers repentant bandits cows for AK-47s . *BBC News - Africa*.

Boudreau, E. P. (2020, March 13). *Supporting One Another in Times of Crisis.* Retrieved from Harvard Graduate School of Education: https://www.gse.harvard.edu/news/uk/20/03/supporting-one-another-times-crisis

Charles L., R. (2002). The Definition of Terrorism. *Analyses of Social Issues and Public Policy*, 9-14.

CNN Library. (2019). *Boko Haram Fast Facts.* CNN World . Retrieved December 7, 2019, from https://www.cnn.com/2014/06/09/world/boko-haram-fast-facts/index.html

Daniel, G. A. (2005, April). Counterterrorism: A Game-Theoretic Analysis. *The Journal of Conflict Resolution, 49*(2), 183 - 200.

Dr. Chukwurah, D. C. (2015, May). Implication of Boko Haram Terrorism on Northern Nigeria. *Mediterranean Journal of Social Sciences, 6*(3), 371-379.

Fleming, J. (2005). 'Working together' : Neighborhood Watch, reassurance policing and the potential of partnerships. *Australian Institute of Criminology*.

Harvard Health Publishing. (2020, May 13). *Preventing the spread of the coronavirus* . Retrieved from Harvard Medical School : https://www.health.harvard.edu/diseases-and-conditions/preventing-the-spread-of-the-coronavirus

Institute for Economics & Peace. (2019, November). *Global Terrorism Index 2019: Measuring the Impact of Terrorism.* Retrieved from http://visionofhumanity.org/reports

Laqueur, W. (1987). *The Age of Terrorism* . Boston : Little Brown.

Ochoche, J. M. (2013, November). Modeling Boko Haram: A Game Theoretical Approach. *International Journal of Science and Technology, 2*(11), 777-783.

Office of the Director of National Intelligence. (2020). *Joint Duty* . Retrieved from Office of the Director of National Intelligence: https://www.dni.gov/index.php/who-we-are

Passas, N. (2005). *INFORMAL VALUE TRANSFER SYSTEMS, TERRORISM AND MONEY LAUNDERING.* Boston: U.S. Department of Justice.

Richardson, Harry W., Gordon, Peter, Moore, James Elliott. (2007). *The Economic Costs and Consequences of Terrorism.* Cheltenham, UK: Edward Elgar Publishing.

Shan, X. and Zhuang, J. (2014, March). Modeling Credible Retaliation Threats in Deterring the Smuggling of Nuclear Weapons Using Partial Inspection - A Three-Stage Game. *Decision Analysis, 11*(1), 44.

Shan, X., Zhuang, J. (2014, March). Modeling Credible Retaliation Threats in Deterring the Smuggling of Nuclear Weapons Using Partial Inspection - A Three-Stage Game. *Decision Analysis, 11*(1), 44.

START. (2009-2019). *Global Terrorism Database*. Retrieved from National Consortium for the Study of Terrorism and Responses to Terrorism:

https://www.start.umd.edu/gtd/search/Results.aspx?search=IDP+Nigeria&sa.x=10&sa.y=14

United Nations High Commissioner for Refugees. (2001-2020). *Nigeria Emergency*. Retrieved from UNHCR: https://www.unhcr.org/en-us/nigeria-emergency.html

Watson, D. L. (2002, February 06). *Before the Senate Select Committee on Intelligence* . Retrieved from The Federal Bureau of Investigation: https://archives.fbi.gov/archives/news/testimony/the-terrorist-threat-confronting-the-united-states

Woodhead, J. (1990). The Role of the Community in Crime Prevention - Neighbourhood Watch and the broader Community Interests. *Australian Institute of Criminology - THE POLICE AND THE COMMUNITY IN THE 1990s*, 84-85.

Zhuang, J., & Bier, V. M. (2007). Balancing terrorism and natural disasters—Defensive strategy with endogenous attacker effort. *Operations Research*,55(5), 976-991.

The Case for Integrating Lean Management with Pharmacy Automation

ABSTRACT

Introduction/Purpose/Background

Two business elements impacting pharmacy operations today are lean management and automation. The purpose of this paper is to analyze the advantages of integrating lean management ('lean') with automation, and propose the optimal mixture of lean and automation, based on pharmacy size and type of automation.

Lean is widely used in manufacturing, but less commonly found in pharmacies. Automation varies by pharmacy size and volume, but even small pharmacies frequently use some basic automated equipment.

Methods

A combination of literature review and the authors' professional observations and experience are used to create the proposed model and recommendations. The analysis considers four distinct types of automation technology, as well as four different categories of pharmacy operations, defined by size and market.

Results

The analysis includes a plot of automation versus lean level for current state versus future state, segmented by pharmacy size. In addition, a table is presented summarizing the effectiveness of lean versus automation in pharmacy operations, based on various performance metrics.

Although pharmacies have increasingly automated the prescription filling process, they often overlook lean management methods as a necessary and complimentary part of their business.

Conclusion

Pharmacy automation seems largely driven by the desire to save labor and mistake proof the medication dispensing process. All pharmacies can benefit from lean in order to reduce waste, but pharmacy size should drive automation levels. Regardless, pharmacies should always analyze process flows from a lean perspective prior to automation, to avoid 'automating the waste'.

Keywords: Lean management; Pharmacy automation; Six Sigma; Process improvement

INTRODUCTION

Lean and Six Sigma

The phrase "lean production" was coined in the late 1980's by John Krafcik, a researcher working at the Massachusetts Institute of Technology (MIT) with the International Motor Vehicle Program, to describe the Toyota Production System.(1) The Toyota Production System (TPS) is heavily centered around continuous quality improvement and waste reduction, production triggered by customer demand and process simplification.(2)

Activities such as reviewing bank loan applications, processing insurance claims, and treating medical patients also involve process steps that can be improved with lean methods.(3) Therefore, because of its applicability to service processes, lean production is often referred to as "lean management", or frequently "lean".

In general, lean focuses on reducing waste, while Six Sigma focuses on reducing process variation using a team-oriented approach requiring estimates of financial benefits. Six Sigma systems traditionally deploy project teams improving processes using a structured methodology for solving problems known as the DMAIC framework; Define-Measure-Analyze-Improve-Control.(4) The DMAIC model was subsequently adopted by companies wishing to integrate lean methods with Six Sigma, and therefore 'lean Six Sigma' programs are predominantly driven by this same problem-solving approach.

Today, lean is typically integrated with Six Sigma principles, and referred to as "Lean Six Sigma" or simply "Lean Sigma". Henceforth, the use of the term 'lean' in this paper includes Six Sigma tools and methods as well.

Lean approaches are used extensively in manufacturing, but are not as widely found in healthcare. Some important lean ideas, such as mistake proofing, have been used in healthcare

for years, but often without formal identification. For instance, the simple idea of systematically tracking instruments during surgical procedures represents mistake proofing, which is a key component of lean methods.(5)

A misconception regarding lean management is that it is a disguised method for reducing staffing. This misunderstanding possibly stems from the phrase 'lean and mean' popular during the 1980s. However, laying-off employees is counterproductive since knowledgeable workers are lost, and any remaining employees have less enthusiasm for improving processes. Therefore, lean companies realize that process improvements should not result in layoffs, until the workforce has been reduced through natural attrition, and attempts to reassign or re-train employees have been depleted.(2)

Recent actions by large companies demonstrate the continued association of the words 'lean' and 'layoff'. For example, in July 2022, Renton, Washington-based Providence announced plans to reorganize and layoff workers in order to create a 'leaner' operating structure. The company announcement stated that "Administrative support departments will adapt their structure to support the leaner operating model".(6) Equating lean with layoffs is also the erroneous view held by many nursing unions. This impedes buy-in when legitimate programs intended to improve efficiency and reduce waste are proposed, because they are considered by some a direct threat to staffing levels.

Lean in Healthcare

Overall, healthcare lags behind the production industry in terms of lean deployment and maturity, so manufacturing remains an important benchmark for healthcare. The quality guru Joseph Juran suggested that when implementing lean methods in healthcare, the experience of other industries should be considered. Juran argued that since managerial processes are alike for

all industries, the principles of lean production could be applied in healthcare, despite being originally developed for manufacturing applications.(7)

Indeed, today there are many examples of lean principles being applied in healthcare. For example, Kaiser Permanente's East Bay location, is an advocate for lean methods and is currently implementing these principles throughout its regional area. The Kaiser Permanente Diablo location uses localized terminology such as "Diablo GPS" in lieu of lean within its Patient Care Improvement System. The Kaiser Permanente South Sacramento Medical Center successfully used lean methods in its emergency department to decrease key performance metrics such as patient length of stay, and door-to-doctor time.(5)

Although the previous examples illustrate the application of lean management in healthcare, they also demonstrate how it can be unevenly emphasized at large providers such as Kaiser Permanente. Lean programs are not standardized across the entire Kaiser Permanente system, and are often siloed to specific regions based on management knowledge of quality tools, and the interests and motivation of local leadership.

Lean in Pharmacies

Pharmacies essentially behave like production processes, with work-in-process inventory consisting of various medications, bottles, vials, and paper labels, and filled prescriptions analogous to the finished goods found in manufacturing. Many articles and books have documented successful applications of lean management in manufacturing, but there are also articles describing the use of lean principles to improve pharmacy operations. Many of these studies include successful applications of lean methods in both community and hospital pharmacies.

Hospital pharmacies are high-cost departments typically representing 10-20 percent of the average hospital's operating budget, and are critical to quality patient care.[8] Pharmacies are therefore under pressure to reduce expenses in order to improve the hospital's bottom line, without affecting quality of care. For example, the Charleston Area Medical Center (CAMC) Health System in West Virginia reduced waste by introducing smaller batch sizes for its sterile intravenous (IV) medication. Initially, IV products were prepared once daily and then delivered to the nursing units for administration to patients. However, this approach created excessive waste in the form of expired and unused medications. The pharmacy staff decided to reduce batch sizes and switch to five deliveries per day, and this simple change resulted in $134,000 annual savings in IV drug waste.[9]

Some pharmacy cost reductions, while being small individually, can collectively add-up to significant savings. For example, Indiana University Health - Bloomington Hospital adopted 50 lean cost-saving recommendations for its pharmacy that, taken together, saved the pharmacy $775,000, with $375,000 coming in the first year.[8] As another example, Cape Cod Hospital in Hyannis saved more than $100,000 annually by not administering high-priced brand name anesthetics. The hospital instead used sevoflurane, a lower cost generic, as a replacement for most of its patients.[8]

The role of the pharmacist is changing, moving away from the stereotypical hospital pharmacist working in the basement. Pharmacists are now found making patient rounds with physicians, and assisting with drug therapy recommendations.[8] Part of the reason for this switch to higher value-added work for pharmacists is the result of myriad small lean improvements, as well as the use of automated pharmacy systems.

PHARMACY AUTOMATION

Authors and clinicians have advocated for the use of automated drug-dispensing devices since the early 1980's, when the first generation pill counting devices were introduced. Yet, automation technology does not appear to be in widespread use. This lack of adoption is partly due to the high cost of automation. For example, a robotic dispensing system can cost more than $500,000, and require another $100,000 annually for support, maintenance, and upgrades. This large expense can be justified only for busy pharmacy locations, filling thousands of prescriptions daily, where automation provides significant labor savings. Small, neighborhood pharmacies might fill 500 prescriptions on a busy day. Therefore, even though automation can be very appealing to smaller pharmacies, the high upfront cost generally cannot be justified. Shingo describes the notion of "automation with a human touch", referring to it as "autonomation".[10] Autonomation requires 'intelligent' machines which can detect an error as it occurs, and take steps to shut down the process. A smart machine would also then take corrective action to prevent the mistake from happening again. Therefore, the ideal pharmacy robot would dispense drugs to every patient with 100% accuracy, in terms of quantity or dosage, and correct medication. A smart robot would shut down automatically if an error were detected anywhere in the process, and corrective action would be taken by the machine. This capability in automated systems is generally not available today, so human intervention and oversight are still required.

In manufacturing applications, automation can reduce direct labor cost and increase productivity. When automation is used in pharmacies, in addition to labor savings, it can reduce dispensing errors and mistake-proof processes, thus improving customer satisfaction. Some common benefits include; ensuring correct pill counts, dispensing the right medication, reducing labeling

errors, reducing order lead time, and reducing inventory storage space. However, benefits can vary by type of automation. For example, a simple pill count machine does not guarantee that the correct medication is dispensed, nor does it necessarily reduce inventory space required for the pill containers.

It is beneficial at this point to clarify the meaning of the term "pharmacy automation", because it includes a wide variety of technologies and layouts. The specific automation type is often referred to as a 'product line' by automation equipment providers. The analysis that follows is restricted to the following four automation technologies:

a. *Blister packaging machines (BP)*. These allow pharmacies to create customized dosage packs for patients. At small pharmacies, blister packing is often performed by hand, but the process can be automated for larger businesses. Blister packs can reduce product tampering and provide good visual control of proper dosing for both the pharmacy and the patient.

b. *Automated storage & retrieval (AS/RS) systems.* These systems typically retrieve native containers from storage and transport them to the pharmacist or technician. The remainder of prescription-filling is completed by hand, transferring pills into proper bottles or vials. The AS/RS technology found in large hospital pharmacies or mail order pharmacies is analogous to the automated systems used to locate and move items in distribution centers and libraries. For large volume operations with huge warehouses, such as Express Scripts, this type of automation is essential. Strictly speaking, from a lean perspective, no value is added with this technology, when only material storage and transport take place. However, smaller pharmacies use systems such as the RxSafe as a non-value added necessity for pharmaceutical compliance and regulatory means.

c. *Robotic dispensing systems (RDS)*. These systems are considered "turnkey" because they can locate, count, fill, and label prescription vials automatically. The pharmacist or technician only verifies labeling, and medication shape and color before dispensing to the customer. This system can be coupled with an AS/RS system to further streamline the process. Robotic dispensing systems are typically limited to large pharmacies where the volume is sufficient to justify the equipment cost.

   Also included in this category is the simple pill and tablet counting device, a less capable variant of the RDS. These machines improve counting speed and accuracy when compared with manual scales or tray-and-spatula methods. However, as is true for RDS systems in general, the user must ensure that the correct medication type is loaded into the hopper.

d. *Mobile Automated Dispensing Cabinets (MADC)*. These systems are primarily used in hospital and patient care locations. They are ideal for centralized pharmacy settings in need of multi-patient or location dispensing and management. The cabinets are generally located on patient floors in a locked, designated room. When a patient order is entered into a monitor mounted to the cabinet, the correct drawer and bin unlock, allowing access to the required medication. A pharmacy technician travels to the hospital floor and stocks the cabinets as needed; the cabinets generally do not return to the central pharmacy.

It should be noted that variants exist for these automation types, including highly technical systems which integrate multiple categories of automation into a single workflow. For example, computer-controlled drug storage devices or cabinets and vial filling capabilities often coexist within a single automation machine. These machines allow for increased control and tracking of

drug distribution and storage. Each machine has a different function, but all have value-adding benefits. In pharmacy fulfillment centers, such as central fill and mail order sites, the end product (i.e., the customer order) varies considerably, and the formulary can be very large.(11)

INTEGRATING LEAN AND AUTOMATION

Common Lean Business Models

The authors' research uncovered situations in the pharmacy industry, where the phrases *quality improvement* and *working more efficiently* are sometimes used when referring to cost reduction.[a] This alternative naming of lean approaches can be observed in manufacturing as well, and it is not uncommon for entities to create terms for process improvement programs that are rooted in lean or Six Sigma methods. Sometimes companies create entire business models derived from well-known existing systems. For example, the *Toyota Production System* (TPS) was rebranded by the Danaher Corporation and Ford Motor Company, and named the *Danaher Business System* (DBS) and *Ford Production System* (FPS), respectively.(12,13) These business models were largely based on Toyota's lean approach to automobile manufacturing.

Labor Considerations

Toyota has long promoted using automation for unpleasant or dangerous tasks, but it also realizes some tasks are performed better by humans. Likewise, when considering pharmacy automation, management should first carefully consider which work elements pose risks or dangers to stakeholders. These dangers would include risks to patients posed by dispensing incorrect medication, expired medication, or the wrong dosage. Arguably, the most important stakeholder in the pharmacy value stream is the patient. Pharmacies must be careful to avoid simply "automating the waste", because an underlying inefficient process cannot be easily fixed

simply through automation. Therefore, prior to installing automated machinery in a pharmacy, management focus needs to be on process improvement using lean principles.

Using the basic equation; productivity = output/input, pharmacy automation has the potential to decrease labor input, thus increasing productivity. However, in the pharmacy, the output needs to be the correct medication and dosage. The four main types of automation technology described earlier have their capability and limits. For instance, the reliability of any AS/RS or automated dispensing cabinet is only as good as the quality of the raw material input (in many cases, capsules or pills). The output quality depends upon several factors, including the pharmacy technician filling bins or pill cassettes with the correct product. In other words, technology helps to reduce prescription errors and allows pharmacists more time to work with patients, but it is only as good as the underlying system - technology will not fix an inherently bad process. This is the same principle that defines software output in the field of computer science; 'garbage in, garbage out' (GIGO).

Automation can reduce labor input from a variety staff positions, including technicians and pharmacists. Decreasing the time pharmacists spend with prescription-related work is a significant benefit, particularly in a hospital setting, where pharmacists would then have more time to visit patients and provide medication advice to doctors.

PROPOSED FRAMEWORK

Four different types of pharmacy operations, categorized by size and available resources, are defined as follows:

Category 4 - Large mail-order facilities

Category 3 - Large hospital pharmacies or clinics

Category 2 - Pharmacies having multiple locations with corporate ownership (e.g., CVS or Walgreens)

Category 1 - Small family-owned pharmacies, or those with limited locations

Consider these four pharmacy categories at their current states, and proposed or future states, as shown in Figure 1. The proposed states suggest the proper mix of lean methods and automation. Facilities with low "lean level" lack organized improvement programs and have minimal workforce engagement in process improvement. They might use simple visual labeling systems to identify and locate inventory, and display basic work instructions, but they lack coordinated improvement efforts or in-depth understanding of lean principles. A highly lean facility would demonstrate advanced lean methodologies, including dedication to continuous improvement and project-oriented practices following the DMAIC approach described earlier. In such facilities, lean training would be required for all members of the workforce, and a certification infrastructure would exist for lean project leaders.

The vertical axis in Figure 1 represents the level of automation in the facility or location. Hand counting and manual prescription filling methods lie at the lower end of the scale. Moving vertically, with increasing automation levels, would be found the usage of blister packaging machines, simple pill and tablet counting devices, automated dispensing cabinets (where appropriate, in a clinical setting such as a hospital), automated storage and retrieval systems, and robotic dispensing systems.

All four types of pharmacy operations are initially placed at the lower half of the lean scale, based on the authors' field observations and survey of the literature. Automation levels generally vary with size of the facility, which is correlated with prescription volume. Figure 1 suggests that all operations should increase lean level going forward, and that all categories, with

the exception of Category 4, tend to underutilize automation. Category 4 facilities have high automation levels initially, usually AS/RS systems, due to their extremely high production volume. Therefore, the current and proposed states for Category 4 facilities are placed at approximately the same automation level. Category 2 facilities initially have moderate to good knowledge of lean principles because of corporate-wide continuous improvement programs. Category 1 locations are placed at the lowest lean level initially, despite the fact that employees are often directed to "work more efficiently", as was the case at the Patterson Family Pharmacy, noted earlier. However, regardless of the obvious need to be more efficient, the workforce at Category 1 locations usually lacks the training or tools necessary to support an effective and sustained process improvement initiative. Based on our field observations, small pharmacies that install automated equipment, before they study and improve processes with lean methods, can be masking further gains in quality and productivity. For example, using basic lean tools such as the 5S's (Sort, Straighten, Shine, Standardize, Sustain) to improve general workplace organization prior to installing automated equipment, can greatly benefit pharmacy productivity. Lean should be used as a starting point for process improvement because it is low-cost approach that can be adopted by any pharmacy, regardless of size. Indeed, without spending any money on automation, low cost lean techniques to improve pharmacy operations are available. For example, the pharmacy department at the University of North Carolina hospitals successfully promoted a culture of continuous improvement using lean Six Sigma as a tool. It used kaizen events to improve a variety of processes, including: code trays, returned medications, office supply inventory management, and triage technicians. The university did not need to purchase automation for any of these projects.[14]

To summarize, Figure 1 suggests that in the ideal state, the proper mix of automation and lean methods can benefit all pharmacy operations, regardless of size. In general, all pharmacies need to maximize their leanness level and adopt automation technology appropriate for their size and prescription volume. Category 1 operations, due to their small size, would not need to create a formalized infrastructure for lean certification or project tracking and reporting.

Some of the more important process performance metrics targeted within a pharmacy setting are listed in Figure 2. The figure also includes the related lean waste or cost of quality category for each targeted metric. Lean management splits waste into eight categories, often remembered with the word "DOWNTIME"; Defects, Overproduction, Waiting, Not using people very well, Transportation, Inventory, Movement, and Excess processing. Each of these are considered non value-added waste elements.

The effectiveness of lean methods and automation at addressing and improving the performance metrics is shown in the two remaining columns. Note that the identifiers were limited to three levels, *Good, Fair*, and *Poor*. Also, the identifiers were assigned considering lean and automation separately, this is, by judging the effectiveness of lean without automation, and automation without lean. This comparison highlights some interesting differences between lean and automation, but also highlights their complementary nature.

First, in our judgement, lean methods are better at ensuring that the correct medication is dispensed. Automating prescription filling works accurately only if the medication is in the correct bins or cassettes, and a source of error exists when relying exclusively on automation due to the GIGO principle. Simple mistake proofing methods upstream would ensure that the bins are filled correctly in the first place. Therefore, best practice uses both lean and automation to reduce prescription errors.

Both lean and automation are effective at reducing pharmacy labor (at all levels, technicians and pharmacists). However, automation accomplishes this directly by physically replacing people labor with machine time, resulting in significant labor reduction. Lean can reduce labor by studying a process, mapping out the workflow, identifying the non value-added steps, collecting and analyzing data, and suggesting improvements. Therefore, using lean to reduce labor can take time, requiring detailed knowledge of the process, data collection, patience and perseverance.

Lean methods are helpful in reducing the amount of unused medications. Recall the Charleston Area Medical Center (CAMC) Health System and its use of smaller batch sizes for sterile intravenous (IV) medication. This example demonstrates the usefulness of lean thinking; waste was targeted, and a solution was conceived, without the need for purchasing automated equipment.

Repetitive strain injuries can be directly reduced by automating pharmacies, because machines are perfect for locating and transporting medications, counting pills, and filling and capping vials. These are great benefits for pharmacies, leading to fewer workers' compensation claims and lost-time injuries. It also relieves the fatigue and boredom of repetitive high-volume work. Pharmacy technicians assigned to operate the automated equipment would require training, enhancing their marketable skills and possibly improving job satisfaction. Lean approaches can be used to reduce injuries as well, but only after the process has been studied first, and the DMAIC process applied. Therefore, the effectiveness was rated as 'fair' for lean and 'good' for automation.

Automation is superior at reducing theft of controlled substances. For example, the mobile automated dispensing cabinets used in clinics and hospitals use locking features to prevent

product loss and ensure the chain of custody. The manufacturer of the dispensing cabinet incorporates lean mistake proofing principles into its design, such as colored lighting on drawers to indicate which medication to retrieve. Mistake *minimization* is possible in this situation, but not mistake *proofing* because the process still requires a human to visually ensure that the proper medication has been placed into the correct unlocked drawer.

The final two rows in Figure 2 highlight a key distinction between lean and automation. Lean is focused on simplicity and waste reduction and carefully examines each process for improvement. Duplicate steps, reports, procedures, or tasks are carefully studied and eliminated. Without using lean methods prior to installing robotic machinery and computers, the pharmacy runs the risk of *automating an underlying inefficient process*. Lean methods are ideally suited for examining duplicative reports and complex administrative procedures for simplification or elimination.

Upon further examination of Figure 2, it is important to note a few additional summary points. First, there are often direct and immediate benefits of automation, whereas lean can take time and is most effective as a continuous improvement program, incorporating small, incremental process improvements. However, a powerful aspect of lean is that it focuses on simple, low-cost approaches that do not require capital expenditures or complicated machines. Second, many lean principles are critical elements of automated equipment and are woven into machine design, such as mistake proofing using colored lights, physical locks, audible warning sirens, etc.

Despite lean emphasizing simple, low-cost solutions to process improvement, lean methods are not opposed to using technology for reducing process waste. For example, consider the elimination of paper drug charts in favor of electronic medication administration records (eMAR) to serve as the legal record of the drugs administered to a patient. The eMAR is a very efficient method for sharing medication information between healthcare providers because it

reduces waste, including several of the eight waste categories explained earlier. However, a lean project following the DMAIC process would not typically result in recommendations to purchase the hardware and software required for an eMAR system. In other words, the lean process emphasizes developing low-cost solutions for improvement, rather than spending money on automated equipment or IT solutions. Regardless, it is hard for lean practitioners to argue against the use of eMar, since its use reduces waste and minimizes errors.

CONCLUSION AND FUTURE WORK

This paper presents a first attempt at determining the optimal mix of automation and lean management, and it was proposed that pharmacy size largely determines this optimal mix. All pharmacy operations, regardless of size, should use at least basic lean tools in order to reduce waste. However, in general, only larger pharmacies should create highly structured lean programs, with project teams following the DMAIC process, and where certification plans are in place for lean facilitators. Not all pharmacies benefit from automation, and the value of automation largely depends upon prescription volume and the type of automation being considered. However, going forward, as equipment prices fall, smaller pharmacies may find it more practical and cost effective to purchase automation.

An interesting and useful future study would involve collecting sample data from pharmacies with varying levels of automation and lean, in order to quantify any productivity differences. This work would first require objective metrics for assessing the 'leanness' of a pharmacy. Most current methods that measure 'lean level', such as Rapid Plant Assessment (RSA)[15], are designed for manufacturing facilities, whereas an approach specifically customized for the pharmacy setting would be needed. Measuring automation level is more straightforward, largely because a casual tour of any pharmacy would provide a good indicator of automation level.

Productivity could be estimated in terms of (prescriptions per day) / (labor hours per day), or an equivalent metric that is readily developed from an interview with pharmacy management. Several unanswered questions could be addressed by collecting this data. For example, when pharmacies install automation, have they first considered lean methods for process improvement, and which has more impact on pharmacy efficiency and productivity, lean methods or automation?

FOOTNOTES

a. On-site interview with Linsey McMinn, PharmD, Patterson Family Pharmacy, Patterson, CA, November 13, 2017. One of the systems that the authors observed at this site was the RxSafe 1800 robotic dispensing system.

REFERENCES

1. Marchwinski C, Shook J, Jones DT. Lean Lexicon – A Graphical Glossary for Lean Thinkers, 4th ed. Cambridge: The Lean Enterprise Institute; 2008.

2. Arnheiter E. Teaching Lean Management using Experiential Methods to Improve Learning. International Journal of Information and Operations Management Education. 2014; 5(4): 344-362.

3. Arnheiter E, Maleyeff, J. The Integration of Lean Management and Six Sigma. The TQM Magazine. 2005; 17(1): 5-18.

4. Pyzdek T, Keller P. The Six Sigma Handbook: A Complete Guide for Green Belts, Black Belts, and Managers at all Levels, 3rd. Ed. New York: McGraw Hill; 2009.

5. Arnheiter E. Kaiser Permanente: Creating a no-wait emergency department. Ivey Publishing. 2016; 9b16m173; 14 pages.

6. Muoio D. Providence targets 'leaner' operating structure, leadership layoffs to offset recent losses. fiercehealthcare.com. 2022; July 20 (accessed online 2023 Jan 13). https://www.fiercehealthcare.com/providers/providence-targets-leaner-operating-structure-leadership-layoffs-offset-recent-losses.

7. Manos A, Sattler M, Alukal G. Make healthcare lean. Quality Progress. 2006; 39(7): 24.

8. Edwards R. In struggle to cut expenses, hospitals eye the pharmacy. Hospitals and Health Networks. 2011; Nov: 28-32.

9. Traynor K. I.V. batching improvements cited in Baldrige win. American Journal of Health-System Pharmacy. 2016; 73(3): 104.

10. Shingo S. A study of the Toyota Production System. Portland: Productivity Press; 1989.

11. Innovation Associates; Central Fill/Mail Order. https://innovat.com/central-fillmail-order/ (accessed 2018 June 14)

12. Henry Ford Health System; Henry Ford Production System. https://www.henryford.com/hcp/academic/pathology/production-system (accessed 2023 Jan 13)

13. Danaher Business System. http://www.danaher.com/how-we-work/danaher-business-system (accessed 2023 Jan 13).

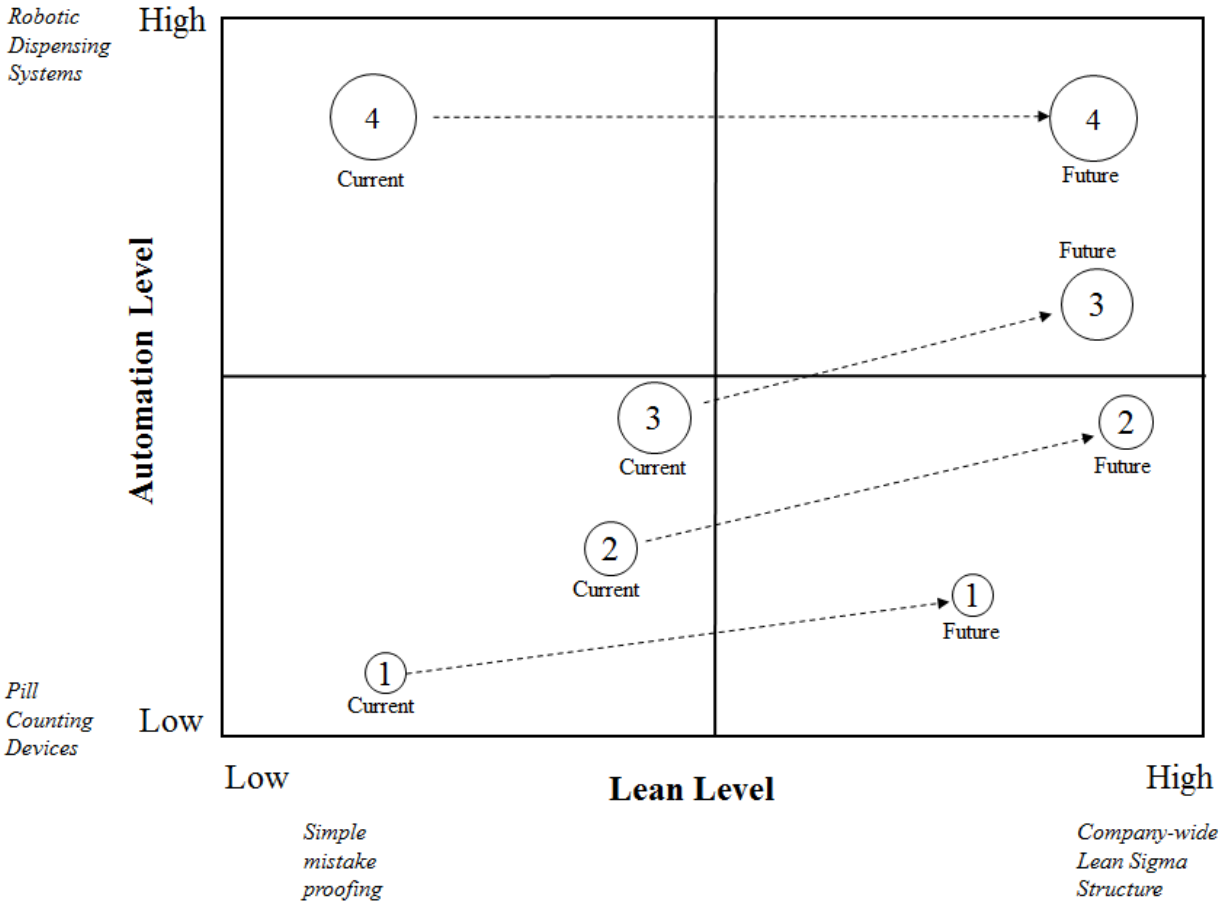
14. Mann J, Dipper A, Rowe E, Hatfield C. Lean Six Sigma in the Pharmacy Department. Pharmacy Times. 2014; September 17 (accessed online 2023 Jan 13). https://www.pharmacytimes.com/publications/health-system-edition/2014/september2014/lean-six-sigma-in-the-pharmacy-department.

15. Goodson RE. Read a Plant - Fast. Harvard Business Review. 2002; 80(5).

FIGURE CAPTIONS

Figure 1. Current State and Future State, by Pharmacy Size (based on prescription volume)

Figure 2. Effectiveness of Lean versus Automation in Pharmacy Operations


Robotic Dispensing Systems
High
Automation Level
Pill Counting Devices
Low
4
Current
4
Future
Future
3
3
Current
2
Future
2
Current
1
Future
1
Current
Low
Lean Level
High
Simple mistake proofing
Company-wide Lean Sigma Structure

Performance Metric Targeted (important within a pharmacy setting)	**Waste or Cost of Quality Addressed (NVA Activities)**	**Effectiveness at Improving Performance Metrics**	
		Lean Methods	Automation
Correct medication is dispensed	Defects (Toyota calls "rework")	Good	Fair
Reduce pharmacy labor	Not using people very well	Good	Good
Reduce time between prescriptions	Changeover time	Fair	Good
Reduce amount of unused meds	Overproduction and inventory in excess	Good	Poor
Reduce repetitive strain injuries (RSI)	Workers' compensation claims	Fair	Good
Reduce variability between patient orders	Excessive process variation	Good	Good
Improve inventory control	Inventory losses	Fair	Good
Reduce theft of controlled substances	Inventory losses and internal failure costs	Fair	Good
Reduce wasted motion of pharmacy staff	Motion in excess	Good	Good
Reduce order lead time for patient	Waiting & Transportation (in clinical setting)	Fair	Good

Reduce required pharmacy floor space	Inventory	Fair	Good
Facilitate pull production discipline	Excess inventory	Good	Fair
Reduce need for complex machinery or complex administrative procedures to perform simple tasks	Excess processing	Good	Poor
Reduce duplicative reports, documentation or data entry	Excess processing	Good	Poor

Sustainability Management

Analysis of the Impact of Climate Change on Hospital Services in Boston

John Maleyeff, maleyeff@bu.edu
Xiaomeng Niu, xniu9@bu.edu
Yujue Tan, cntanyj@bu.edu
Mengya Zhao, ameliazh@bu.edu
David Weidman, dweidman@bu.edu

Boston University
Metropolitan College

Abstract

An analysis was undertaken to explore the impacts of climate change on the hospitals in Boston. The focus was on supply chain impacts that included both demand side changes as well as potential changes in the ability of hospitals to service patients caused by future supply chain disruptions. Methodologies included a meta-analysis of prior work that would be relevant to the hospitals in Boston and a data mining effort to quantify some of the anticipated impacts. Results indicate the potential for a number of impacts on patient demand, including increases in the frequency of patients with existing illnesses and the potential introduction of new maladies. Regarding their supply chains, the low altitude coastal location of many Boston hospitals will lead to disruptions in last mile delivery and the potential exists for additional power outages that will disrupt many of the outsourced medical treatment services.

1. Introduction

Climate change will significantly affect the weather in Boston (Massachusetts), a coastal city with wide temperature extremes and substantial precipitation. According to Douglas and Kirshen (2021), by 2050 the average temperature in the Northeast U.S. is projected to rise by 2.2°C (reduced emissions scenario) or by 2.8°C (business-as-usual scenario). Their report projects that the Boston area will see daily maximum temperatures above 32°C on about 28 days by 2050 and about 48 days by 2100. Projections also include a 20-30% increase in precipitation intensity by 2100. Under the most optimistic scenario, by 2100 the sea level in Boston harbor is projected to rise between 35 to 78 cm. All of these factors will have a significant impact on Boston's public health, energy demand, agriculture, natural resources, infrastructure, and transportation.

These climate change trends will impact Boston citizens' lifestyle, including their health. The hospitals in Boston will likely need to accommodate their patients by adjusting healthcare practices based on new or increasing patient demand, while dealing with more disruptions in their ability to serve those patients. More persistent high temperatures, extreme weather, flooding, and rising sea levels will likely exacerbate health problems for residents, including the introduction of new viruses and diseases. Flooding, power failures, and other disruptions will create supply chain challenges for hospitals by affecting employee availability, supply logistics, and outsourced services that require reliably accessible and available technologies.

The Greater Boston healthcare infrastructure includes 25 hospitals and 20 community health centers, with the Longwood Medical Area representing a world-class health epicenter. There are more than 46,000 scientists, researchers, and staff, and over 21,000 students studying at institutions and hospitals (boston.gov). The healthcare sector is growing in concert with the Greater Boston region's population, which grew by almost 1,000,000 residents in the ten

years preceding 2020 (from its 2011 population of 4.9 million). According to the Massachusetts Department of Public Health (mass.gov), the demand for hospital and health care systems just prior to the COVID-19 pandemic numbered 4.7 million annual visits a year. Hospitals outsource or offshore a significant portion of their activities, including the manufacture of medical devices, medical supplies, and pharmaceuticals, as well as many services including administrative services and medical services (such as MRI imaging analysis).

This article reports preliminary research to identify and assess the risk factors brought on by climate change that would impact Boston hospitals' supply chain. It begins the process of quantifying those risk factors to determine the magnitude of their impacts. This methodology considers impacts on hospital supply chains, both patient demand and the hospitals' ability to service their patients. It is a methodology that should be applicable to hospitals in other regions, although the details need to be location specific. First, a representative meta-analysis of prior research is employed that considers the public health effects that are relevant to the Boston region. Particular attention is paid to research that reported quantifiable impacts that can be used to project corresponding impacts in Boston. Second, original data mining analyses were performed using publicly available data sets. The main data sources were obtained from the Harvard dataverse (dataverse.harvard.edu), an energy data warehouse (data from eversource.com), the National Drought Mitigation Center (droughtmonitor.unl.edu), and the U.S. National Oceanic and Atmospheric Administration (NOAA.gov). The main analysis tool was predictive modeling using multiple linear regression.

2. Meta-Analysis: Public Health and Supply Chain Impacts

A plethora of research concerning climate change impacts on public health has been published. Here, a representative sample of published work is presented to illustrate the range of climate change effects. Many of these impacts have already been noticed and are expected to become more prevalent in the next decades. For example, 58% of all known pathogenic diseases have already been exacerbated by climate change (Mora et al, 2022). Many researchers have addressed how climate change will affect public health around the world, including an estimated increase of 250,000 deaths annually between 2030 and 2050 (Hayes et al, 2018). Rising temperatures will increase the prevalence of a heat-related illness (HRI), such as heat cramps, heat exhaustion, and heat strokes (Khan, 2019; Fuhrmann, 2016). Yin et al (2019) determined that high heat exposure on 4 or more consecutive days increased HRI-related emergency department (ED) visits significantly. Van der Linden et al (2019) found that ED visits increased significantly for children during hot days.

Humidity increases are also expected to impact the incidence of HRI's (Sarmiento, 2016). People who work outdoors will be especially affected because high humidity makes it more difficult for the body to dissipate heat (Spector, 2019). Mosquito-friendly habitats will expand, causing malaria, dengue fever, chikungunya, and West Nile virus to become more common in previously unexpected regions (Jordan, 2019). In Massachusetts, tick-borne diseases are also projected to increase (McDermott, 2022; Massachusetts DPH, 2017). Romanello et al (2021) found that the Boston region can see increases in various Vibrio bacteria-related diseases such as gastroenteritis and sepsis.

Storm frequency and intensity increases have been shown to cause higher levels of vehicle collisions (Eisenberg 2004). Dust particles associated with sandstorms are known to cause lung diseases as well as effects on the heart, skin, and brain (Ardon-Dryer et al, 2019). Floods combined with abnormally warm sea surface temperatures favor cholera outbreaks (Caminade, 2019). Wildfire occurrences were shown to increase cardiac arrest and the

occurrence of cardiac disease (Wettstein, 2018). Waterborne disease outbreaks are preceded by heavy precipitation events in the U.S. (Curriero et al, 2001) and the U.K. (Nichols et al, 2009), and extreme precipitation was linked to waterborne infections in Taiwan (Chen et al, 2012). Mental health will also be affected. Hayes et al (2018) cited evidence that extreme weather events can trigger post-traumatic stress disorder, depression, and anxiety.

Today's pharmaceutical supply chains are complex, with active ingredients, secondary processing, and labeling rarely co-located, and distribution networks that extend across the globe (Halim et al, 2019). The COVID-19 pandemic caused many delays in non-vaccine related clinical trials due to transportation disruptions (Weidman et al, 2021). Climate change is having and will continue to have a well-known impact on these supply chains (Song et al, 2022). Pharmaceutical supply chain disruptions due to climate change have already started, and they are especially disruptive when extreme weather events happen in quick succession (Levermann, 2019).

The biggest impact of climate change to transportation systems is reduced stability of road or other logistics infrastructures (Strauch, 2015). Extreme weather can raise sea levels that result in disruptions for all transportation modes - marine, land-based, and air transportation (EPA, 2008). Rail tracks have deformed when the air temperatures have been extremely hot (McGuirk, 2009). Sea level rise will cause flooding that will temporarily or permanently damage airports, roads, rail lines and tunnel (USGCRP, 2009). For example, California's Highway 1 collapsed, Michigan experienced unprecedented flooding, and Texas saw rare snowstorms (U.S. DOT, 2022). Flooding caused worldwide medical supply shortages during 2021 floods at the Port in Malaysia (Munir et al, 2022).

A recent trend in supply chains concerns the "last mile" logistics systems that are expected to grow by 36% from 2020-2030 in the world's 100 largest cities (Joselow, 2020), with logistics costs around 50% in many cases (Gopal & de Miquel, 2017). This trend will be particularly important to Boston's hospitals, because many of them are located in known flood plains. It will impact patients' ability to travel to a hospital and it will impact the delivery of items both inbound and outbound. These impacts may be particularly problematic for pharmaceuticals and other perishable products that may require cold storage during shipment (Nadimuthu & Victor, 2022). They can also affect items being shipped out of hospitals, such as blood or tissue samples (Gutierrez-Franco, 2021), and can be particularly impactful when disruptions occur in underserved urban communities (Jones et al, 2021).

Hospitals rely on electricity to serve patients, ensure information security, operate equipment, and communicate with outsourced service providers. But electricity generation becomes less efficient as the ambient air temperature increases. By 2050, estimates include a 9% increased demand (Veselka et al, 2007) and a 10-20% increase in the need for additional electric generating capacity (Wilbanks et al, 2008). Extreme weather occurrence can damage energy supply facilities, as evidenced by Hurricanes Katrina and Rita that damaged more than 100 platforms and damaged 558 pipelines in 2005, impacting markets as far away as New York and New England (USGCRP, 2009).

Higher temperatures and associated humidity will make workers more susceptible to illnesses and they will be more likely to make errors (Parsons, 2014). As such, workers may need to take more frequent breaks (USGCRP, 2016). According to Albert et al (2021), outdoor high-intensity workers will experience more problems that may require hospital visits and hospital personnel productivity will decline. In some regions, the average annual labor productivity for high intensity work will decline in the next decades (Knittel et al, 2020). The increased risk of heat-related diseases will result in lower overall labor capacity and wages, even in

indoor spaces (USGCRP, 2016). The market cost associated with the lost labor is estimated to be between $5 and $40 billion per year by 2050 (Zhang & Shindell, 2021).

2.1 Meta-Analysis Results

Figure 1 summarizes the meta-analysis results found in the literature. The left side of the figure lists climate impacts, with color coding used to show where relationships between climate changes and Boston hospital-related impacts are anticipated. The right-side of the figure lists hospital supply chain characteristics (both demand and supply related), with color coding used to illustrate relationships between the hospital-related impacts and a hospital's supply chain. The impacts projected include increases in patient demand, and more frequent disruptions in a hospital's ability to service their patients.

Some of these impacts documented in Figure 1 may be mild and others more extreme. Certain impacts will have a relatively small effect on Boston hospitals. However, their cumulative impacts is likely to be large and it is clear that hospitals should make climate change a major priority. A survey of three well-known Boston hospitals showed a considerable range in their climate adaptation and mitigation efforts. One hospital's climate mitigation plan was robust, while limited planning has taken place at another hospital. The third hospital fell within these two extremes.

3. Data Mining: Quantifiable Impacts

Although the quantitative aspects of this work have just begun, some insights into climate change impacts on Boston's hospitals appear to be quantifiable. Several challenges existed, however, in the determination of specific quantifiable impacts of climate change on these supply chains. First, although weather data are readily available and easily downloaded, historical trends may differ from future climate patterns. Second, detailed response data from specific hospitals (e.g., admissions, morbidity statistics, etc.) are generally unavailable to the public. Third, identifying relationships that apply to Boston based on trends shown to exist in other regions needs to be done with caution.

Historical weather data were obtained from NOAA, which is the umbrella organization that includes the U.S. National Weather Service. The input variables corresponded to conditions relevant to Boston (now or in the future). Downloads of historical NOAA weather data are available to the public, free of charge. These data are provided on an hourly and daily bases by zip code. Many of the weather variables are correlated, such as those concerning various forms of temperatures and precipitation. For daily values, many variables included the minimum, maximum, and average for each day listed. The input variables for the data mining analyses consisted of the following:

1. Temperature (dew point, dry bulb, wet bulb, deviation from normal, wind chill).
2. Barometric pressure (sea level, station).
3. Wind (sustained speed, peak speed, direction).
4. Precipitation (rain, snowfall rate, snow depth, sleet).
5. Relative humidity.
6. Degree days (heating, cooling).
7. Weather description code.
8. Others (fog and smoke index, lightning, flooding, hurricane).

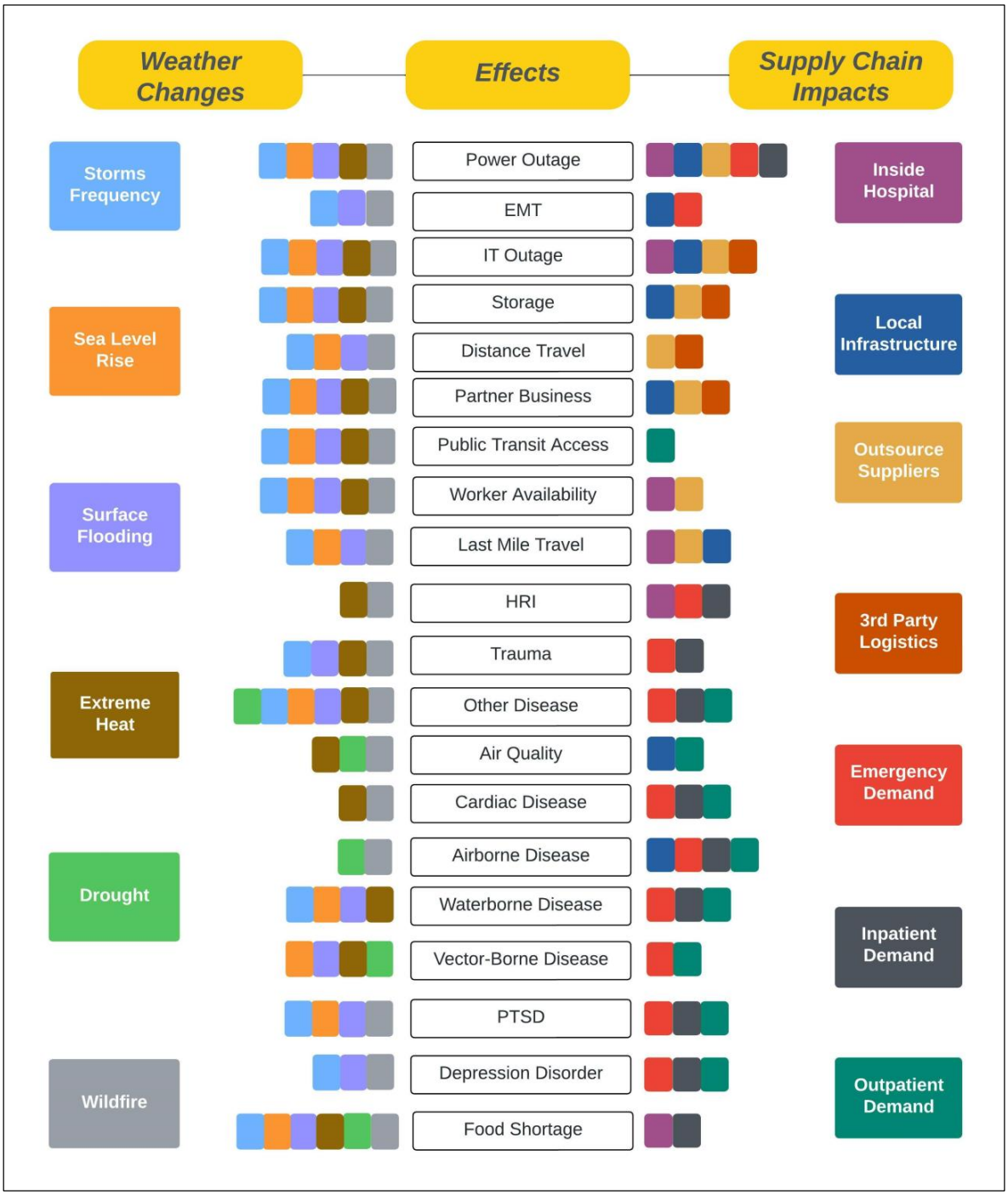


Figure 1: Climate Change Impacts on Boston Hospitals

NOAA also provides data on the number of extreme event days by location. They include a daily indicator for the existence of any of the following: high temperature, high sustained or peak wind, drought (a period of extraordinary dry weather with water shortages causing hydrologic imbalances), blizzard, thunderstorms, frost or freezing conditions (that would harm pipes and sprinkler systems), and flooding.

The response data sets were used to analyze the supply chain impacts that may become apparent in Boston as the 21st century progresses. Although some of the responses are not

directly associated with the Boston region, they can be used as surrogates to predict changes in patient demand or the likelihood of a supply chain disruption. For example, if a climate effect increases ED demand in one part the U.S., that same effect can be expected to increase ED demand in Boston. Several data sources were used to generate response variables, as described below:

1. Data from the study done by Van der Linden et al (2019) were used as response data to analyze effects of extreme weather on ED visits in Sacramento (Van der Linden, 2019). These data sets were collected from 2009 to 2016, and their relationship to NOAA weather events for the Sacramento zip code were evaluated. Although hospitals in Sacramento differ in size from Boston hospitals, the significant weather effects and subsequent tendencies would also be expected in Boston albeit at varying magnitudes.
2. A data set concerning 911 calls in Montgomery County Pennsylvania was used to determine the relationship between extreme weather and certain injuries or diseases (Chirico, 2016). The data were collected from 2015 to 2016. Although these data were collected in Pennsylvania, tendencies found in the relationship between emergency calls and climate would also be expected in Boston albeit at varying magnitudes.
3. A data set concerning power outages in the Boston region was purchased from the Eversource energy provider (the source of these data was https://poweroutage.us/). Eversource is the largest energy services provider in the region, and these data covered October 2017 to January 2021. They included daily numbers of customers without power for any portion of each day, as well as the total number of hours the outages consumed per day.

3.1 Data Mining Results

Table 1 provides quantifiable details of the data mining analysis results, based on multiple regression analyses.

Table 1: Data Analysis Results

Impact	Response Variable(s)	Prediction Models (for all predictor variables, p<0.05)
Emergencies	911 emergency calls in Montgomery County PA (2015-16)	$Y_F = 1.65 + 0.74X_H$ $Y_H = 0.11 + 1.75X_H$
ED Demand	ED Visits in Sacramento, CA (2009-16)	$Y_E = 0.30 + 1.09X_D + 2.36X_F + 2.12X_W$ $Y_E = 196.2 + 5.02X_H$
Power Disruptions	Outages in Boston (2017-21)	$Y_O = 438 + 111X_H + 648X_F + 863X_W + 612X_T$ $Y_C = 1.6 + 4.2X_H + 2.9X_F + 3.7X_W + 3.5X_T$

The following notation is used:

$X_H = 1$ if high temperatures existed (0 otherwise)

$X_D = 1$ if drought was declared (0 otherwise)

$X_F = 1$ if flooding occurred (0 otherwise)

$X_W = 1$ if extreme winds were present (0 otherwise)

$X_T = 1$ if thunderstorms occurred (0 otherwise)

The thresholds for high temperature days varied according to the location of the response data. In Sacramento, the threshold was high temperature at least 32ºC. This threshold was reached on 23% of days. In Boston, the threshold corresponded to a high temperature of 30.5ºC was reached on 4.1% of days. In Pennsylvania, the threshold corresponded to a high temperature of 32.2ºC was reached on 7.6% of days.

The following notation was used for the response variables. Although the baseline demand for ED visits for Sacramento and 911 calls in Montgomery County are not consistent with those in Boston, the models are used to predict the proportional increase for these responses rather than their magnitude.

$Y_H =$ 911 calls for heat exhaustion per day

$Y_F =$ 911 calls for fever per day

$Y_E =$ ED visits per day

$Y_O =$ Number of power outages per day

$Y_C =$ Number of customers (000) impacted by a power outage per day

4. Future Work

As a consequence of this preliminary work, the authors received a grant to focus attention on climate change impacts to the Boston Medical Center (BMC) patient demand and BMC's ability to provide clinical and support services. This work will commence in January 2023 and will include more comprehensive, timely, and specific data on hospital impacts. The project team will identify the climate-related risks and quantify their anticipated impacts on patient demand and service supply activities at the BMC, then offer a set of mitigation strategies. The methodologies will combine the synthesis of prior research relevant to Boston and the BMC that includes risk assessment, data mining and meta-analysis, as well as analyses of data obtained concerning patient demand and hospital supply chains. Mitigation approaches will be explored in conjunction with BMC leadership and technical staff.

5. References

Albert, C., Bustos, P., & Ponticelli, J. (2021). *The Effects of Climate Change on Labor and Capital Reallocation*. Working paper doi: 10.3386/w28995, National Bureau of Economic Research, Cambridge, Massachusetts.

Ardon-Dryer, K., Mock, C., Reyes, J., & Lahav, G. (2019). The effect of dust storm particles on single human lung cancer cells. *Environmental Research, 181*, 108891.

Caminade, C., McIntyre, K.M., & Jones, A.E. (2019). Impact of recent and future climate change on vector-borne diseases. *Annals of the New York Academy of Sciences, 1436*(1), 157-173.

Chen, A., & Murthy, V. (2019). *How health systems are meeting the challenge of climate change*. Harvard Business Review. Retrieved November 17, 2022, from https://hbr.org/2019/09/how-health-systems-are-meeting-the-challenge-of-climate-change

Chirico, M. (2016). Emergency - 911 Calls: Montgomery County, PA, Harvard Dataverse, Version 1, doi: 10.7910/DVN/X8Q4YA, UNF:6:343G1vJlwHxmqBqsSsPbzQ==.

Curriero, PF.C., Patz, J.A., Rose, J.B., & Lele, S. (2001). The Association Between Extreme Precipitation and Waterborne Disease Outbreaks in the United States, 1948-1994. *American Journal of Public Health, 91*(8), 1194–1199.

Douglas, E., & Kirshen, P. (2021). *Climate Change Impacts and Projections for the Greater Boston Area: Findings of the Greater Boston Research Advisory Group Report*. University of Massachusetts Boston, Boston, Massachusetts.

Eisenberg, D. (2004). The mixed effects of precipitation on traffic crashes. *Accident Analysis and Prevention, 36*(4), 637–647.

EPA (2008). *Climate Impacts on Transportation*. U.S. Environmental Protection Agency, Washington, DC.

Fuhrmann, C.M., Sugg, M. M., Konrad, C. E., & Waller, A. (2016). Impact of extreme heat events on emergency department visits in North Carolina (2007–2011). *Journal of Community Health, 41*(1), 146–156.

Gopal, G., & de Miguel, A. (2017). Tackling the difficult last mile. *Industrial and Systems Engineering at Work, 49*(12), 28–33.

Gutierrez-Franco, E., Mejia-Argueta, C., & Rabelo, L. (2021). Data-driven methodology to support long-lasting logistics and decision making for urban last-mile operations. *Sustainability (Basel, Switzerland), 13*(11), 6230.

Halim, I., Ang, P., & Adhitya, A. (2019). A decision support framework and system for design of sustainable pharmaceutical supply chain network. *Clean Technologies and Environmental Policy, 21*(2), 431-446.

Hayes, K., Blashki, G., Wiseman, J., Burke, S. & Reifels, L. (2018). Climate Change and Mental Health: Risks, Impacts and Priority Actions. *International Journal of Mental Health Systems, 12*(1), 2840.

Jones, E.C., Azeem, G., Jones, E. C., Jefferson, F., Henry, M., Abolmaali, S., & Sparks, J. (2021). Understanding the last mile transportation concept impacting underserved global communities to save lives during COVID-19 pandemic. *Frontiers in Future Transportation, 2*.

Jordan, R. (2019). How does climate change affect disease? *Stanford Earth Matters Magazine*. Stanford University, Stanford, California.

Joselow, M. (2020), Delivery vehicles increasingly choke cities with pollution. *Scientific American, E&E News*, https://www.eenews.net/.

Khan, A.A. (2019). Heat related illnesses. Review of an ongoing challenge. *Saudi Medical Journal, 40*(12), 1195–1201.

Knittel, N., Jury, M. W., Bednar-Friedl, B., Bachner, G., & Steiner, A. K. (2020). A global analysis of heat-related labour productivity losses under climate change-implications for Germany's foreign trade. *Climatic Change, 160*, 251-269.

Levermann, A. (2014). Make supply chains climate smart. *Nature (London), 506*(7486), 27-29.

Massachusetts DPH (2017). *Massachusetts State Health Assessment*, Massachusetts Department of Public Health, Boston, Massachusetts.

McDermott, A. (2022). Climate change hastens disease spread across the globe. *Proceedings of the National Academy of Sciences - PNAS, 119*(7), e2200481119.

McGuirk, M., Shuford, S., Peterson, T.C., & Pisano, P. (2009). Weather and Climate Change Implications for Surface Transportation in the USA. Word Meteorological Organization Bulletin n°: 58(2).

Mora, C. *et al.* (2022). Over half of known human pathogenic diseases can be aggravated by climate change. *National Climate Change, 12*(9), 869–875.

Munir, Z.A., Sundram, V. P. K., & Adham, K. A. (2022). Inter-Organisation Practices for Humanitarian Supply Chain: A Case Study of Flood Disaster in Malaysia. *Global Business and Management Research, 14*(1), 194.

Nadimuthu, L.P.R., & Victor, K. (2022). Environmentally friendly micro cold storage for last-mile Covid-19 vaccine logistics. *Environmental Science and Pollution Research International, 29*(16), 23767–23778.

Nichols, G., Lane, C., Asgari, N., Verlander, N.Q., & Charlett, A. (2009). Rainfall and outbreaks of drinking water related disease and in England and Wales, *Journal of Water and Health, 7*(1), 1-8.

Parsons K. (2014). *Human Thermal Environments: The Effects Of Hot, Moderate, And Cold Environments On Human Health, Comfort, And Performance*. CRC Press, Boca Raton, Florida.

Romanello, M., et al (2021). The 2021 report of the Lancet Countdown on health and climate change: code red for a healthy future. *The Lancet (British Edition), 398*(10311), 1619–1662.

Sarmiento, S. (2016). It's not just the heat, it really is the humidity: Know the risks. NBCNews.com. NBC News, New York.

Song, Z., He, S., Wang, Y., & An, J. (2022). Green pharmaceutical supply chain coordination considering green investment, green logistics, and government intervention. *Environmental Science and Pollution Research International, 29*(42), 63321–63343.

Spector, J.T., Masuda, Y.J., Wolff, N.H., Calkins, M., & Seixas, N. (2019). Heat exposure and occupational injuries: Review of the literature and implications. *Current Environmental Health Reports, 6*(4), 286–296.

Strauch, R.L. (2015) Adapting transportation to climate change on federal lands in Washington State, U.S.A. *Climatic Change, 130*, 185-199.

U.S. DOT (2022). *Climate Action*. U.S. Department of Transportation, Washington, DC.

USGCRP (2009). Global Climate Change Impacts in the United States. In Karl, T.R., J.M. Melillo, and T.C. Peterson (eds.), *Climate Change Impacts by Sectors: Energy Supply and Use*, U.S. Global Change Research Program. Cambridge University Press, New York.

USGCRP (2016). The Impacts of Climate Change on Human Health in the United States: A Scientific Assessment. Crimmins, A., et al (eds.). U.S. Global Change Research Program, Washington, DC.

Van der Linden, N., Longden, T., Richards, J. R., Khursheed, M., Goddijn, W. M. T., van Veelen, M. J., Khan, U. R., & van der Linden, M. C. (2019). The use of an 'acclimatisation' heatwave measure to compare temperature-related demand for emergency services in Australia, Botswana, Netherlands, Pakistan, and USA. *PLoS One, 14*(3), e0214242.

Van der Linden, N. (2019). ED visits and temperature, Harvard Dataverse, Version 1, doi: 10.7910/DVN/QHPZOX.

Veselka, T.D., Botterud, A., Conzelmann, G., Koritarov, V., Poch, L.A., & Wang, J. (2007). Climate Change Impacts on the Electric Power System in the Western United States. *Eos, 88*(52).

Weidman, D.A., Zang, Y., Maleyeff, J., Corlu, C.G., and Song, Y. (2021). Impacts of COVID-19 on the Work Environment of Professional Workers: Implications for Educators & Human Resource Managers, Research Report 2021-01, BU MET Decision Sciences Research Laboratory.

Wettstein, Z. S., Hoshiko, S., Fahimi, J., Harrison, R. J., Cascio, W. E., & Rappold, A. G. (2018). Cardiovascular and cerebrovascular emergency department visits associated with wildfire smoke exposure in California in 2015. *Journal of the American Heart Association, 7*(8).

Wilbanks, T.J., et al (2008). *Effects of Climate Change on Energy Production and Use in the United States*. U.S. Climate Change Science Program: Synthesis and Assessment Product 4.5, Washington, D.C.

Yin, Q., Wang, J. (2017). The association between consecutive days' heat wave and cardiovascular disease mortality in Beijing, China. *BMC Public Health, 17,* 223.

Zhang, Y., & Shindell, D. T. (2021). Costs from labor losses due to extreme heat in the USA attributable to climate change. *Climatic Change, 164*(3-4).

Impacts of ESG on Gen Z and Gen Y Employees' Prosocial Motivation in the Workplace

ABSTRACT

The prosocial behaviors of quality talent increase profitability, productivity, and efficiency and sustain competitive advantages**.** As a result of attracting talent, retaining talent, and other retention challenges, business leaders continue to implement various strategies to mitigate prosocial motivators and improve quality of work. This research examines environmental, social, and corporate governance ESG and its ability to impact employee prosocial motivators, as a whole and by generation (Gen Z versus Gen Y), through the strategic and targeted implementation of business initiatives.

From a self-administered online questionnaire via Amazon Mechanical Turk and students at a public university in the New England, 845 responses are used to test our hypotheses via multiple regression and correlation analysis.

Employee prosocial behaviors varies based on the type of ESG initiatives businesses adopt. Environmental ESG has the highest impact on prosocial motivation, followed by social ESG. Additionally, employee generation affects which ESG should be prioritized. For Gen Y, environmental ESG has the highest impact on prosocial motivation, followed by social ESG. For Gen Z, social and then environmental ESG influence prosocial motivation.

This study is among the first to examine how different ESG initiatives impact prosocial motivation behaviors and employee generation moderates the connection.

ESG initiatives can benefit businesses that are looking to engage in employee prosocial behaviors and improve business-specific factors that lead to increased profitability, innovation, as well as employee satisfaction and engagement.

These study results increase empirical evidence with management literature and may contribute to promoting ESG initiatives in the business community.

Keywords:

Prosocial motivation; Environmental, social, corporate governance; ESG; Employee generation, Regression analysis

1. INTRODUCTION

Liao et al. (2022) reported that prosocially motivated employees have higher job satisfaction, perform better compared to their peers, and advance further in their careers. Garrett et al. (2021), employee productivity can be predicted by prosocial motivation. Those with high prosocial motivation show a more negative response to self-interest, meaning decision-making that exhibits behavior misaligned with progressive organizational goals will result in negative responses for those with high prosocial motivation, and an indifference to those with low prosocial motivation. Zhu et al. (2014) reported that individuals with high prosocial motivation when engaging in helping behavior, are strongly affected by social influences and are less concerned with personal rewards and consequences, whereas those will low prosocial motivation help others due to self-interest and expectations of personal returns. More precisely, Bakker, et al (2015) speculate that prosocial motivation strengthens the relationship between job resources, work engagement, and increased performance, and weakens the negative relationship between job demands, exhaustion, and performance. Individuals react differently depending on their level of prosocial motivation. Those individuals with low prosocial motivation will perform their duties with the intent of receiving individual recognition and benefit versus the benefit of others. Individuals that thrive from prosocial motivation behaviors are likely to create resources as compared to depleting resources at work Spreitzer et al. (2012). A thriving workforce itself is a competitive advantage as it helps in avoiding negative individual outcomes like stress, depression, and work burnout and promotes positive organizational outcomes high performance, job satisfaction, and organizational loyalty. Additionally, individuals that thrive from prosocial motivation at work are less likely to show absenteeism and stress. When considering employee conduct within an organization, prosocial behavior is defined as any activity which an employee demonstrates for the benefit of others when performing job duties Shin, et al (2020).

An organization that strives to improve the welfare of others external to that organization as related to the social factor in Environment, Social, and Governance (ESG), has employees which perceive that they are working in an ethical climate and exhibit prosocial motivation (Shin et al., 2020). Yang and Basile (2022) studied the effects that CSR plays on the Top 100 global brands. They found that CSR has a positive impact on a firm's performance. They also found that this holds true when external stakeholders become involved in the initiatives. In the grouping of individuals by age, generations are used because of shared similarities and characteristics among those of the same generation. These shared similarities stem from mutual experiences on a historical, social, and cultural level Mahmoud et al (2021). As the older baby boomer generation retires from the workforce, priorities on environmental, social, and governance initiatives will change as values from younger generations become more essential and widespread throughout organizational decision-making. A company that is ESG-compliant serves their potential customers and partners to view the company's credibility and commitment to values such as sustainability and diversity. Which will also serve as a catalyst for new ways to do business and an inspiration for new offerings and services, such as green versions of established products (Fantini et al., 2022).

This research aims to consider the impact of ESG factors affecting employee motivation in the workplace. This paper also wondered if there were any differences between two employee generations (Gen Y and Gen Z) in the workplace, in terms of the impacts of ESG on work motivation. More specifically, this study wondered (1) if environmental factors had any impacts

on workplace motivation, (2) if social factors had any impacts on workplace motivation (3) if corporate governance factors had any impacts on workplace motivation.

Much research has investigated the positive impacts of environmental factors on employee prosocial motivation (Hur et al., 2018; Chen et al., 2020; Suganthi, et al., 2019; Sipahutar et al., 2022; Amy Yeo Chu May et al., 2021). There have been numerous studies on the positive relationship between social factors on work engagement (Chien et al., 2020; Suganthi et al., 2019; Kajackaite et al., 2020; Skudiene and Auruskeviciene et al., 2010) Many articles have reported empirical evidence of positive impacts of corporate governance on employee engagement (Stazyk, Davis, and Liang 2021; Shakil, H. 2020; Damayanti et al. 2021; Gladas and Piderit 2009). However, few studies have been done on how all three factors – environmental, social, and corporate governance together affect employees' prosocial motivation in the workplace. Furthermore, few studies examined the differences of such impacts of the three factors among two generations – Gen Y and Gen Z employees. Additionally, no article has reported the generational differences in the impacts of the three ESG variables on employees' prosocial motivation in the workplace.

Therefore, this research attempted to answer the research question. This paper conducted an empirical study and investigated the generation differences of the impacts of the three ESG factors on the employees' prosocial motivation. The results would fill the gap in the management literature.

This paper developed three main research hypotheses based on the literature and created a survey instrument in line with the hypotheses. This research collected sample data from Amazon Mechanical Turk in 2020 and the students in the public university in New England area from 2021 to 2022. To test the hypotheses, this study ran a statistical analysis of the survey data using regression analysis.

In Chapter 2, this research provides a review of prior studies as a basis for the development of the research hypotheses. Chapter 3 presents the research methodology with the statistical results of hypothesis testing following in Chapter 4. Chapter 5 discusses the results of the study as compared with the literature and articulates the managerial implications. Finally, Chapter 6 concludes this study.

2. LITERATURE REVIEW

Impacts of ESG on Prosocial Motivation/Work Motivation

Hur et al., (2018) examined service employees' perceptions of CSR and how it affects their creativity, compassion, and intrinsic motivation at work. To test the hypothesis of employees' perceptions of CSR and employee creativity sequentially influences compassion at work employees' intrinsic motivation data was collected from hotel employees in South Korea. A total of 480 frontline employees in four full-service luxury hotels were invited to complete 255 questionnaires, and a total of 250 questionnaires were obtained. This study revealed that both employees' compassion and intrinsic motivation play an essential role as a mediating link between employees' CSR perceptions and employee creativity. Therefore, managers should understand that employees are very sensitive to their firms' CSR activities, which can either promote or deteriorate their intrinsic motivation through compassion at work.

Shin et al. (2020) conducted a study in which the team anticipated that service employees' perceptions of their organizations' social responsibility activities would encourage them to become prosocially motivated. The team hypothesized that perceived CSR would be positively related to prosocial motivation. To test the hypotheses of this study, they collected data from front-line service employees of luxury hotels that were engaged in CSR activities such as greening, energy efficiency efforts, water conservation, climate change awareness campaigns, and community donations. The participants were asked to complete a paper-and-pencil survey. Of the 650 research questionnaires, 435 survey questionnaires responded. The findings of this study indicate that the prosocial activities of a service organization spread to organizational members who are aware of them, in turn having a positive impact on the employee's level of prosocial motivation and behavior.

Chien et al. (2020) studied different influences on work motivation within the hospitality industry in Mongolia. 398 questionnaires were filled out. Multiple hypotheses were formed, notably that the enjoyment of work has significant positive effects on employee performance, and internal self-concept has significant positive effects on employee performance. An important finding in the results was the significant positive relationship between work motivation and the organization's commitment to individual performance and goal internalization. There is a direct correlation between organizations being supportive of employee goals, values, and beliefs, and higher workplace motivation and performance. There is significant evidence that employees desire to work for an organization that enables them to use their own skills to realize their own personal potential.

Lanfranchi and Narcy (2022) try to understand the effectiveness of pay for performance among employees that have a high prosocial motivational drive. The study looks at the effectiveness of performance-rated pay in both the non-profit and for-profit sectors. It is shown that employees in the non-profit sectors are not affected as much by the pay-for-performance strategy as their for-profit counterparts. Employees of the non-profit sector are more motivated by their impact on the environment than the money they earn.

Hur et al., (2018) found that service employees' perceptions of CSR and positively affects their creativity, compassion, intrinsic motivation at work. Chien et al. (2020) found a direct relationship between employee-supportive practices in organizations and workplace motivation and performance. Shin et al (2020), found that the prosocial activities of a service organization spread to organizational members who are aware of them, in turn having a positive impact on the employee's level of prosocial motivation and behavior. Lanfranchi and Nancy

(2022) prove that money is not the only motivator for employees. Prosocial motivation is a huge incentive for employees in the non-profit sector. As a result, there is an advantage for companies to be socially responsible that benefits employees and shareholders.

Impacts of Employee Generations on Prosocial Motivation/Work Motivation

Calk and Patrick et al. (2017) examined the factor influencing Millennials' workplace motivation. The five investigated factors were based on the Workplace Motivation Inventory (WMI) model after Maslow's Hierarchy of Needs and Herzberg's Hygiene-Motivator Model of Satisfaction which measured: basic, safety, belonging, ego-status, and actualization. A total of 341 surveys were randomly distributed among lower and upper-division classes across a liberal arts university in the Southwestern United States. Of the 314 surveys distributed, 88 surveys were returned and completed in their entirety. To test for the importance of each motivational need, the WMI are measured on a scale of 0-100 with a higher score indicating greater importance of the need. A pairwise comparison of WMI shows that Millennials are willing to take career-related risks to experience meaningful and satisfying work as long as they can meet basic needs. Rather than motivated by stable, secure jobs with predictable salaried and benefits based on low scale on the safety motivational need. The evidence of this study proves that Millennials are diverse in their motivators within a workplace.

Wong et al. (2008) analyzed whether personality and motivational driver differences exist across three generations in the workplace: Baby Boomers, Gen Xs, and Gen Ys. They developed two research hypotheses which proposed the following: There would be key differences in personal preferences across the three generations, and secondly that there would be key motivational differences across the three generations. To test for differences across the three generations, an existing dataset of participants' responses to the Occupational Personality Questionnaire and the Motivational Questionnaire were used. A one-way ANOVA was used to test each hypothesis. The results showed Employees from different generations were motivated by the following factors: affiliation, power and progression. On motivational drivers such as immersion, ease and security, and personal growth, no significant differences across generations were found.

Solnet, et al. (2012) studied the differences in the workplace between generation Y and older generations. The study was conducted in the hospitality industry in Australia. It was noted that the hospitality industry provides solid insight into workplace attitude because of the strong emphasis placed on coworker relationships and the role it plays on job satisfaction. The study was conducted using focus groups with generation Y employees and focus groups with hospitality owners and managers. The results showed a strong desire for generation Y workers to have robust friendships of trust within an organization. This generation seeks a strong company culture with laid-out values that emphasize respect. Individual growth is important, so companies should provide incentives for employees to learn and develop.

Einholf (2016) did a comparison between Generation X and Millennials when looking at prosocial motivation. The thought was that younger generations are more likely to be active in their communities and giving back socially. It does not seem to hold true as the two generations are right next to each other. It would make more sense when looking at two generations that were separated quite a bit, such as Baby Boomers and Millennials. It is significant to point out that generations do not necessarily think differently on issues, and they can tend to agree on specific views.

Calk and Patrick et al. (2017) found that Millennial workers have diverse motivating factors within a workplace. Result of the analysis reveals that Generation Y's personal success is significantly affected by servant leaders who display accountability and forgiving aspects rather than modesty and empowerment. Wong et al. (2008) found that on motivational drivers such as immersion, ease and security and personal growth, no significant differences across generations were observed. Solnet, et al. (2012) explored the differing ways in which Generation Y find workplace motivation. The study found that those in Generation Y desire company provided initiatives to further improve their own personal development and education. They also wish the company they work for to have core values that provide them with respect. Einholf (2016) found that there is not much of a difference between Generation X and Millennials when it comes to prosocial motivation. These studies reveal that generationally there are significant different impacts on work motivation based on individual needs.

Impacts of Employee Generations on ESG

Klimkiewicz et al. (2017) conducted research to analyze how do Millennial job seekers attitudes toward corporate social responsibility (CSR) influences employer attractiveness (EA). The data collection was determined by two assumptions. First, investigating employer attractiveness by Millennials who were actively looking for a job. Second, Millennials' job experiences, their job search process, and the importance of corporate social responsibility in it. The concentration of Millennials represented the Y Generation, which are expected to be sensitive toward ethical issues. Data was collected at job fairs, for young job seekers, 339 questionnaires were obtained voluntarily and anonymously. The respondents consisted of undergraduate (44.2%) and graduate students (18.3%). Multivariate analysis was applied to test three independent variables: H1.1, corporate social responsibility normative, H1.2, corporate social responsibility positive, and H1.3, corporate social responsibility engagement. The dependent variable was measured through a self-constructed corporate social responsibility-based EA measurement instrument which consisted of corporate social responsibility importance for employer choice and reject a job offer from an employer who does not fulfill basic corporate social responsibility requirements. Millennial job seekers are likely to both acknowledge corporate social responsibility importance during a job search and reject a job offer if the employer neglects corporate social responsibility.

Alonso-Almedia (2018) performed a study to analyze how millennials perceive companies to determine if this generation specifically is attracted to companies with more sustainable environmental practices. This research proposed a hypothesis in which responsible companies are more attractive to millennials. To test the above hypotheses, in 2014, a structured questionnaire was administered to 1,833 university students from six different countries. An exploratory factor analysis was performed for each factor. A Varimax rotation was then conducted to achieve a one-dimensional structure. The hypothesis that millennials perceive responsible companies as good workplaces, or in other words, millennials are attracted to responsible companies because they could identify with their values was accepted.

Luger et al. (2022) analyzed the support of corporate social responsibility among consumers of different generations. The three hypotheses based on generation stated that Millennials perceived the importance of social responsibility, economic responsibility, and environmental responsibility as important, and all have a positive influence on CSR awareness. Data was collected from both China and Austria, which 441 total participants. 256 responded

from Austria, and 185 from China. Findings showed that generation Y showed views of high importance towards CSR, while millennials showed that CSR was becoming increasingly important to them. Managerial implications in the study state that targeting CSR initiatives should be different for different generations; generation Y has an inherent desire for CSR, while millennials have a learned desire for CSR.

Reavis et al. (2021) look at the comparison between Millennials when it comes to making decisions on either CSR or maximizing profits. The study was a 9-week simulation that required about 50 decisions to be made weekly that would affect shareholder value. The study found that Millennial were more likely to focus on CSR, as they have a responsibility to do good in the world. No other generations were studied; but if there was a comparison, it would show that older generations would be in favor of maximizing shareholder value. This is a good study to show how a generation such as Millennials think, but we could also point to the fact that there were no real-world implications when doing the simulation. Would the scores still be the same if they had pressure from shareholders? In the end, it does show that Millennials at least think about the societal impacts they have.

Klimkiewicz et al. (2017) discovered that Millennials job seekers' attitudes toward corporate social responsibility is highly important in employer attractiveness. Alonso-Almedia (2018) found that Millennials tend to perceive responsible companies as good companies to work at. Reavis et al. (2021) showed that Millennials are more likely to use profits to engage in Corporate Social Responsibility over reinvesting to maximize shareholder value. Luger et al. (2022) reaffirmed that Millennials look at CSR as a major aspect of how a company should operate yearly. Overall, Millennials resonate with companies that are socially responsible.

Environment-related ESG and Prosocial Motivation/Work Motivation

Chen et al., (2020) studied the relationship between environmental stewardship and job satisfaction. To obtain a more comprehensive insight between environmental stewardship related to climate change, emissions, waste management, and job satisfaction a total of 299 questionnaires were collected from five-star green restaurants in the cities of Tainan and Kaohsiung which were accredited by the Environmental Protection Administration in Taiwan. The study found that environmental stewardship positively affects job passion in the green restaurant context. Additionally, there is a positive linkage between job satisfaction and life satisfaction that occurs when a person has a great job passion thus providing support for environmental-friendly concepts. Therefore, employees' environmental stewardship orientation will reflect their willingness to devote more effort to working in a job that promotes environmental issues.

Amy Yeo Chu May et al. (2021) conducted analysis to determine whether an organization's environmentally sustainable activities when communicated to employees within the organization motivate employees. The organization should ensure positive outcomes are communicated from an environmental aspect, by emphasizing the importance of employee participation in support of environmentally sustainable activities. The data was collected via a survey that involved 213 employees currently employed in socially responsible companies with CSR and sustainability reports that are publicly accessible. The data obtained were analyzed using SPSS 25 and Smart-PLS. The study found that employees who work for a company that is environmentally sustainable have a positive effect on employees' level of organizational trust.

Suganthi (2019) examined the relationship between corporate social responsibility initiatives and employee performance. The companies studied had over 100 employees and had been in business for five or more years. The study hypothesized a positive relationship between the adoption of green practices and performance. Data was collected from 580 employees in supervisory roles between April and August of 2018. The study found that CSR initiatives had led to an increase in market performance, followed by performance increases in operations, environmental practices, and cost efficiency. It was indicated that a reason for this relationship stems from the desire of stakeholders to push toward green practices. This motivated employees on an environmental level, as well as an economic level. The theoretical implications of the study show the two-fold benefit of ESG on work motivation, as motivation increases because of both economic benefits and personal prosocial beliefs.

Sipahutar et al. (2022) studied the effects of workers' motivation when working to improve environmental problems. They looked at the fish processing industry; this had an impact on the environment. They tested to see if work motivation was needed to increase environmental sanitation. This would have a positive impact on the environment. They tested 96 respondents; of which, more than half resulted in an increase in work motivation and cleaner production. It holds true that an increase in work motivation can positively affect environmental issues surrounding specific types of work.

Chein et al., (2020) found a positive linkage between environmental stewardship and job satisfaction. Sipahutar et al. (2022) found a positive correlation between increased work motivation and working to fix environmental issues. Suganthi (2019) studied the relationship between green practices and employee performance, stating that green practices improve employee performance because of both environmentally conscious and economic reasons. In conclusion, there is a strong correlation between Environment-related ESG and employee motivation.

Hypothesis 1a: Environment-related ESG is positively related to the employees' prosocial motivation in the workplace.

Generation differences on the impact of Environment-related ESG on Prosocial Motivation (work motivation)

Acheampong et al., (2020) studied six factors that impacted work engagement for Generation Z employees. The identifiable hypothesized factors include CSR, leadership, transformation, work-life balance, autonomy, and technology. A multi-item questionnaire was collected using Amazon Mechanical Turk, with 69 surveys. The results of this study confirm and reinforce that leadership, work-life balance, and autonomy are important factors in employee engagement. CSR, transactional leadership, and technology however showed statistical significance for improving employee engagement.

Day et al. (2013) conducted a study to examine the influence of environmental CSR-related issues in the recruiting process of college-level applicants. Specifically, a perceived relationship between an organization's CSR and its attractiveness as employers to college-level applicants. Environmental programs for consideration included activities like energy conservation, water reduction efforts, waste management, and recycling. The research was conducted through an online survey administered to college recruiters, using a modified snowball approach. The study examined whether there is a perceived relationship between an organization's CSR attracting college-level applicants to hospitality and tourism organizations. Interestingly, "Reputation of the company" and "Opportunity to learn" were the most-important

factors in this list attracting college-level applicants, while "Corporate social responsibility programs" were perceived to be more important than "Salary" in attracting recruits.

Zainee and Puteh (2020) studied the impact of CSR on employee retention, specifically among Generation Y. Notably, the study looked at the ethical responsibilities of companies, which entailed environmental practices. 377 accountants were studied using a questionnaire. It was noted that Gen Y has a sense of responsibility towards the planet, growing up learning the importance of green practices like recycling and using less water. The study proved a significant relationship between ethical (including environmental) responsibility and employee retention among Generation Y employees. The significance of pro-environmental practices for younger generations can be attributed to the environmental teachings which they experienced growing up. Older generations lacked such teachings, thus making environmental CSR practices more important for Gen Y and less important for older generations in work motivation.

McDougle et al. (2011) study the motivation that young adults have on environmental volunteerism. They did a survey at a large urban Canadian university to try and find out why young adults are so keen on volunteering to help the environment and what motivates them to do so. The young generations have lived through so many catastrophic events, that they want to do everything they can to help stop the next one. This shows to carry over to employee motivation when looking for jobs. It shows the motivation of working at a non-profit organization that specializes in environmental issues.

Acheampong et al., (2020) revealed that CSR, transactional leadership, and technology have a statistical significance for improving employee engagement. Day et al. (2013) conducted a study to examine the influence of environmental CSR-related issues in the recruiting process of college-level applicants. Specifically, a perceived relationship between an organization's CSR and its attractiveness as employers to college-level applicants. Zainee and Puteh (2020) found that younger generations tend to embrace environmental CSR practices more so than older generations because of their environmentally conscious upbringing. McDougle et al. (2011) reported the correlation between younger generations and having the motivation to have an environmental impact. In conclusion, generations play an important role when examining the correlation between environment-related ESG and the prosocial motivation of employees.

Hypothesis 1b: Generations moderate the effect of environmental-related ESG on employees' prosocial motivation in the workplace, such that environmental-related ESG is more positively related to prosocial motivation of younger generations.

Impacts of Society-related ESG on Prosocial Motivation/Work Motivation

Society-related ESG as described by the Investopedia Team. (2022) entails a company's impact on the human capital management of its employees and the societal outcome of its decision-making. This factor of ESG pertains to the internal and external relationships sustained by the organization in the community in which it operates, the companies procure to pay supply chain operations, internal and external stakeholders, and employees.

Bernardino et al. (2021), studied the relationship between corporate social responsibility (CSR) and employee engagement. The study aimed to explore if employees and potential candidates want to work for a company that they feel is providing back to society. In other words, employees and potential candidates expect their employeers to be socially responsible and provide tools and resources to engage in causes they care about. The participants were asked to complete a short ten-question five-minute survey on the platform Survey Monkey. The

investigator approached friends and acquaintances through social media and encouraged them to share it with their friends and acquaintances creating a snowball sampling. The survey was also posted on the following social media accounts: Twitter, Facebook, and LinkedIn. 100 responses were collected. The findings of this study indicated that companies whose people are deeply connected to their corporate social responsibility efforts tend to have better workplace motivation than those that do not. Therefore, a company's CSR initiatives and programs are obtaining applicants' interest and can be an advantage for talent acquisition and retention. Moreover, the company will see an increase in work motivation, employee engagement, and opportunities to socially and environmentally be involved within the company's community.

Kajackaite et al. (2020) studies the impact of managers who exhibited prosocial behavior influence on employee workplace motivation. Moreover, selecting a manager with a preference to spend resources on social causes will increase employee workplace motivation. Data for this study was collected using survey responses by 6.557 employees in 947 establishments and contains information about management practices and an employee-level survey on attitudes and personality. The study tested the hypothesis using a regression analysis. Interestingly, having leadership in place which cares for the welfare of others, rather than mere profit maximization will generate trust and motivate employees since their efforts will not be narrowly used only to make as much money as possible for shareholders, and therefore are prosocially motivated.

Frazier and Tupper (2018) studied the effects of the prosocial behavior of supervisors on employees and the resulting psychological impacts. Data was collected from 245 employees and 83 supervisors. The study found a positive relationship between prosocial behavior and employee psychological safety. This is notable as it was stated in the theoretical contributions that prosocially motivated individuals are more likely to engage in proactive behaviors; meaning they are motivated to work and improve their work. The study also found that "employee thriving" had a positive relationship with prosocial motivation. The social aspects of an organization put forth by supervisors show positive aspects on employee behavior. When an employee realizes their supervisor exhibits a genuine interest in their development, a positive relationship occurs with psychological safety. It is imperative for human capital management initiatives to support the training of supervisors to understand the impact their behavior has on employee behavior and motivation.

Skudiene and Auruskeviciene (2010) examine the correlation between CSR and employee motivation. The study was a survey of 274 employees of middle and large corporations in Lithuania. There was a prior study done that showed CSR has a direct impact on turnover, recruitment, etc.. The authors took it a step further to understand if and how CSR would directly affect employee motivation. Their findings show there are two types of CSR motivations: internal and external. It is shown that internal and external CSR activities positively correlate with employee motivation.

Bernardino et al. (2021), found that corporate social responsibility incentives and programs can be an advantage for talent acquisition and retaining workplaces among existing employees. Skudiene and Auruskeviciene (2010) reported that internal and external CSR activities have a positive correlation with employee motivation. Frazier and Tupper (2018) examined the positive relationship between social initiatives by supervisors on the psychological safety of their employees. In conclusion, companies that focus on social-related ESG tend to have more motivated employees.

Hypothesis 2a: Social-related ESG is positively related to employee prosocial motivation in the workplace

Impacts of Generational Employee Work Motivation on Social ESG

Sinha et al. (2022) studied employee work motivation and job satisfaction among Generation Y. The hypothesis indicates that human desires focus on needs starting from the lowest to the highest degree such as physiological needs, stability, social need, self-esteem, and upgrading (Herzberg philosophy). The study focused on many classes between the age of 21-34 years old, and 254 responses were obtained. The results show that exciting jobs, decent working environments, high pay, and promising opportunities are driving factors in Generation Y. Additionally, research reveals that workers in Generation Y tend to do their job, benefiting them in respect to protection and security at work. With a considerable emphasis on making the planet a better place through work, community-focused, and team operations.

Wong et al. (2008) studied motivational differences across three generations in the workplace: Baby Boomers, Gen Xs, and Gen Ys. They developed two research hypotheses which proposed there would be key differences in personal preferences across the three generations, and secondly that there would be key motivational differences across the three generations. To test for differences across the three generations, an existing dataset of participants' responses to the Occupational Personality Questionnaire and the Motivational Questionnaire were used. A one-way ANOVA was used to test each hypothesis. The results showed Employees from different generations were motivated by the following factors: affiliation, power, and progression. Interestingly, regarding motivational drivers such as immersion, ease and security and personal growth, no significant differences across generations were found as per this study.

Wisse et al. (2015) examined the role that employee age plays on job satisfaction in relation to corporate social responsibility. It was noted that working for socially responsible companies positively impacted work engagement and retention. A key aspect of the study was the significance of the perception of future time remaining as an indicator of support for CSR practices. This study hypothesized that as employees get older, limitations in their perspective form, causing a stronger emphasis on emotionally meaningful goals. As a result, this study analyzed data contradicting our own hypothesis, stating CSR desires become stronger as a person ages. Two separate studies were conducted in the Netherlands, one with owners and CEOs from 54 companies, and another with 500 random employees from different companies to support their hypothesis. The study may have benefited from more specific questions, as the context within CSR is very complex, and different results may yield from different CSR-related questions.

Rank and Contreras (2021) were curious to see how generations play a role when it comes to CSR and work motivation. From prior research, it has been proven that there is not strong enough evidence that generations are vastly different when they are sequential (Comparing Gen X to Gen Y). In this study; however, they studied across all generations and were slightly able to see the generational impact that CSR has on employee motivation. Millennials motivation is impacted the most when it comes down to a company's CSR. An interesting takeaway from this study is that the authors think by adding in a cross-cultural element then there would be even more of a correlation.

Sinha et al. (2022) revealed that job satisfaction and work motivation among Generation Y emphasizes on self-fulfillment. Rank and Contreras (2021) reported that younger generations, such as millennials, are more likely to be internally and externally motivated when a company invests in CSR. Wisse et al. (2013) found that employees shift their view on CSR as they age,

becoming increasingly motivated by CSR initiatives. In conclusion, younger generations tend to have more prosocial motivation when their company focuses on social-related ESG.

Hypothesis 2b: Generations moderate the effect of society-related ESG on employees' prosocial motivation in the workplace, such that society-related ESG is more positively related to prosocial motivation of younger generations.

Governance-related ESG and Prosocial Motivation

The last component of ESG as highlighted by the Investopedia Team (2022) is related to the corporate governance practices which a company has implemented. Governance-related ESG considers how the company is led and managed. These practices help ensure a company operates with integrity and diversity, selecting leadership, and providing transparent financial reporting with adequate internal controls. A company that adheres to ESG standards is one that does not use political contributions to obtain preferential treatment or engage in illegal conduct.

Stazyk et al., (2021) examined the relationships between workforce diversity, goals, and job satisfaction. It was assumed that diversity management policies have a direct and indirect effect on the relationships between workforce diversity, goal ambiguity, and job satisfaction. This study hypothesized that both ethnic diversity and diversity management policies directly increase employee satisfaction by clarifying organizational goals. To test the hypotheses data was collected from the Federal Employee Viewpoint Survey (FEVS) over a three-year span (2010, 2011, 2012), the analysis included 1,109,134 federal government employee respondents. Unfortunately, FEVS does not specify racial or ethnic groups but reflected minority or non-minority status. Results support that diversity management policies enhance job satisfaction by clarifying organizational goals. Therefore, employees who believe organizational goals are clear tend to be more satisfied with their jobs.

Shakil, H. (2020) conducted a study to determine the effects of corporate governance mechanisms as related to work motivation and leadership. Corporate governance mechanisms guide long-term strategies of motivation and are significant tools to compete in competitive markets. Compensation and motivation management must be linked according to corporate governance strategy and human resource strategy. The study was conducted via 150 surveys which were used to seek the response of governance mechanisms in the aspects of motivation, leadership, and independent audit committee top-level management and then analyzed using the statistical package for the social sciences (SPSS) via a Pearson correlation of Mean, Standard deviation and regression. This study finds that motivation and leadership mechanisms positively affected employee satisfaction and work motivation.

Damayanti et al. (2021) studied the impact that the implementation of corporate governance and transformational leadership style has on performance by using work motivation as the mediating variable. The study was conducted at a publicly owned hospital with 241 respondents. 84.23% of the respondents were service staff, while the remaining 15.77% were management and administration. The study found significant evidence to support that corporate governance implementation significantly impacts work motivation. More corporate governance initiatives lead to higher work motivation. The main contributions to effect corporate governance implementations were noted as transparency, accountability, responsibility, independence, fairness, and equality.

Gladas and Piderit (2009) examine the effects corporate governance have on employees. There are many aspects when focusing on employees, but a key one is the impact on employee motivation. The study did a survey on American workers that would first capture the level of

corporate citizenship they perceive of their company. After gauging this, the survey can examine many different aspects including their level of motivation. The study found that employee motivation is higher for those that believe their company is serious about corporate governance.

Stazyk et al., (2021) found that workforce diversification and goal clarity directly influences and enhances employees' job satisfaction. Shakil, H. (2020) found that motivation and leadership mechanisms positively affected employee satisfaction and work motivation. Damayanti et al. (2021) found that there is a significant positive relationship between corporate governance implementation and work motivation. Gladas and Piderit (2009) reported that the presence of corporate governance plays a role in directly affecting an employee's motivation. In conclusion, corporate governance plays a positive role in employee motivation and job satisfaction.

Hypothesis 3a: Governance-related ESG is positively related to employee prosocial motivation in the workplace.

Governance-related ESG, Prosocial Motivation, and Employee Generation

Kowske, Rasch, and Wiley et al., (2010) conducted a comparative study of generational differences in workplace attitudes. The data gathered by WorkTrends was compared among Baby Boomers, Generation X, and Millennials. Results reveal that the nature of work such as production to the service sector across all generations are satisfied with their jobs because their skills are being put to good use. However, Millennials have higher job satisfaction when their employer provides recognition, job security, and career advancement compared to Baby Boomers and Generation X. Therefore, this study confirms that Millennials are more satisfied and optimistic than previous generations.

Chan, H. et al., (2016) performed a study to determine the significance between a company's R&D efficiency and CEO employee pay gap. When evaluating the corporate governance factor in ESG, employee compensation and pay gap can be considered. As the researchers found, a pay gap may cause a sense of inequity and cognitive dissonance, which can reduce employees' motivation and productivity. An Employee who perceives a wide pay gap will have reduced work motivation, which in turn may have a harmful effect on the innovative and collaborative behaviors within the R&D teams. All data, including the number of employees and total salary expense, were collected from Compustat. The R&D data is obtained from the RQ database available from Wharton Research Data Services (WRDS).

Made, et al. (2018) studied the personal work values of Generation Y and analyzed the results in relation to other generations. An online survey was conducted in 2018 to people identified as Generation Y. The survey used 56 indicators on 351 respondents using social media. It was stated that while Generation X workers tend to prefer independence in work, Generation Y workers have a higher loyalty to an idea than a company. They also prefer better work-life balance and convenience. The study found that employee engagement is heavily influenced by personal values in Generation Y. It was suggested that companies should use corporate training to improve employee engagement and creativity within Gen Y.

Srivastava and Banerjee (2016) examine the motivations and beliefs of Generation Y as they start to move into the workforce. There was a survey done to find out what will drive the newest generation and how this will affect companies moving forward. There were many different motivating factors that came out of the study, but one key aspect was highlighted. Money is important, but ethics is more important. This highlights the mindset that it is important for their companies to be sound from a governance standpoint.

Kowske, Rasch, and Wiley et al., (2010) revealed that Millennials have a higher level of overall company and job satisfaction, satisfaction with job security, recognition, and career advancements but similar levels of pay satisfaction than previous generations. Made, et al (2018) found that Generation Y is motivated by personal values and requires corporate training to cater to such. Srivastava and Banerjee (2016) found that Generation Y cares about corporate governance more than their own financial well-being. In conclusion, younger generations place a higher emphasis on personal values within their work than older generations.

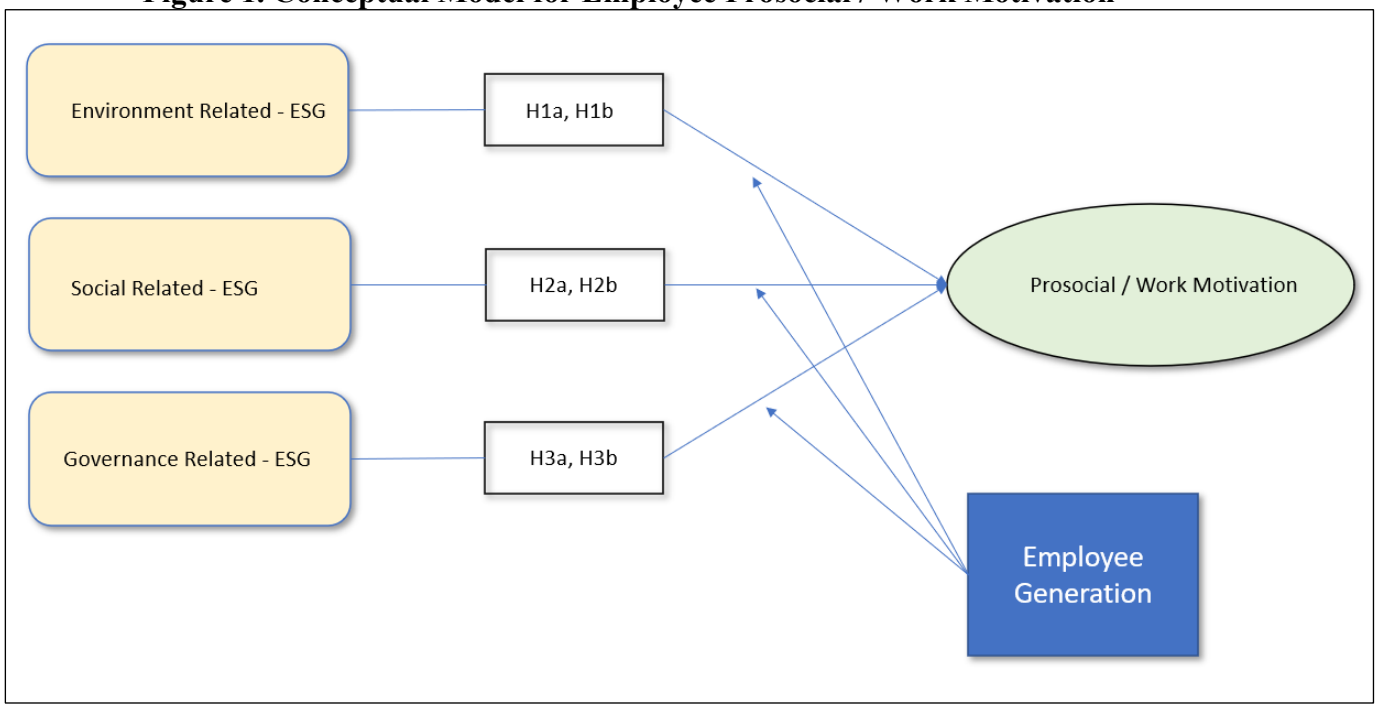
Hypothesis 3b: Generations moderate the effect of governance-related ESG on employees' prosocial motivation in the workplace, such that governance-related ESG is more positively related to Prosocial Work Motivation of younger generations.

3. RESEARCH METHODOLOGY

Conceptual Model

As part of the literature review, three research hypotheses on the impact of ESG on prosocial motivation are proposed. Figure 1 below shows a conceptual model or visual representation of these hypotheses.

Figure 1. Conceptual Model for Employee Prosocial / Work Motivation



Variables

Dependent Variable: Employee Prosocial/Work Motivation **(Y_1)**

Prosocial or work motivation is the dependent variable in this study. Prosocial motivation is important within the work environment as it relates to actions specifically intended to benefit or improve the well-being of other individuals. In other words, the employee is motivated to work harder because the team, or organization, may benefit. This study focuses on three main independent variables which include environmental, social, and corporate governance-related ESG. This study also includes a generation moderator exploring the difference between older generations and newer generations entering the workforce regarding prosocial motivation. This study suggests that ESG factors have a greater effect on younger generations when it comes to employee prosocial motivation.

To assess factors that influence employee prosocial motivation, we selected items from Yu et al. (2021) A questionnaire survey was conducted, and data was collected from 318 respondents who had rich social work experience.

Code	Item	Reference
PRO1	I care about benefitting others through my work	Yu et al. (2021)
PRO2	I want to help others through my work	
PRO3	I want to have a positive impact on others	
PRO4	It is important to me to do good to others through my work	

Grant & Berry (2011) found that when an employee is led by prosocial motivation and considers perspectives other than their own, this will typically lead to intrinsic motivation which produces ideas that are highly creative. Prosocial motivation at its core is the desire to expend effort based on a concern for helping or contributing to other people. This team hypothesized that prosocial motivation would strengthen the association between intrinsic motivation and creativity. The items below were selected for this study.

Code	Item	Reference
PRO5	Because I want to help others through my work	Grant and Berry (2011)
PRO6	Because I want to have a positive impact on others	
PRO7	Because it is important to me to do good for others through my work	

Independent Variables

Woo (2013) highlights a CSR variable with 12 items using five dimensions – Environment, Human Rights & Labor, Product Responsibility, Society, and Economic dimensions. From these dimensions, this study aligns these CSR variables to the three ESG factors for analysis purposes.

***Independent Variable: Environment-related ESG* (X_1)**

Sustainability of the environment has been a key component that many stakeholders want their companies to focus on. Businesses must create and maintain initiatives that proactively try to aid and fix issues within the environment. A survey was done of 377 Gen Y accountants, in which Zainee and Puteh (2020) reported that when companies focus on these environmental activities then employees have a higher motivation to work. Creating a bond between employees and the environment gives employees a sense of purpose that makes them try harder. In this study, there were three items that assessed the environment-related ESG. These items include that a company should produce high-quality products and services that meet society's needs, produce innovative products and services that improve society's well-being and invest in research & development that provides improvement in the value chain.

Code	Item	Reference
	I think the company I work for tries to:	Zainee & Puteh (2020)
ENV1	Produce high quality products and services that meet society's needs	
ENV2	Produce innovative products and services that improve society's well-being	

ENV3	Invest in Research & Development that provides improvement in the value chain	

This paper focuses on 3 of the 6-environment related ESG variables from Woo (2013) These variables focus on a company's impact on the environment and how it influences prosocial motivation. These environmental items are also in alignment with Global Reporting Initiative (2012) The 3 items span across taking care of water, energy, and material uses, minimizing pollution, and investing to protect the environment.

Code	Item	Reference
	I think the company I work for tries to:	Woo (2013) Global Reporting Initiative (2012)
SREN1	Take care of water, energy, and material uses	Environment-related dimension
SREN2	Minimize pollution when producing products/services	Environment-related dimension
SREN3	Invest to protect the Environment	Environment-related dimension

***Independent Variable: Social-related ESG* (X_2)**

Social ESG is concerned with the relationships sustained between the business and its internal and external stakeholders. Social ESG plays a critical role in the perception of a given business and its ability to enhance the lives of those connected to the business. Bernardino et al. (2021), measured the society-related ESG connection between employee engagement through corporate social responsibility programs to measure prosocial motivation. Findings indicated that corporate social responsibility opportunities surge prosocial behaviors.

Code	Item	Reference
SRE1	Creates corporate social responsibility programs	Bernardino et al (2021)
SRE2	Invests in their employees' interest	
SRE3	Provides back to society	
SRE4	Socially and environmentally conscious	

Sinha et al. (2022), measured the social-related ESG connection between age and employee job satisfaction. Findings indicate descent working environments. High pay, and promising opportunities are driving factors in Generation Y.

Code	Item	Reference
SRE5	I care about my basic human desires	Sinha et al (2022)
SRE6	I care about security and high pay	
SRE7	I care about descent working environments	
SRE8	I care about growing opportunities	

Woo (2013) developed a CSR variable with 12 items using the five dimensions – Environment, Human Rights & Labor, Product Responsibility, Society, and Economic dimensions. Among the original 12 items for measuring the CSR variable in Woo (2013), this paper picked two items from the human rights & labor-related dimension (SRH1, SRH3), and tw

items from the product responsibility-related dimension (SRP1, SRP2), one item from the society-related dimension (SRS1), for measuring the Society-related ESG variable as follows:

Code	Item	Reference
	I think the company I work tries to	Woo (2013)
SRH1	Protect human rights at workplaces	Human rights & labor-related dimension
SRH3	Clarify health care benefits for employees	Product responsibility-related dimension
SRP1	Clearly label/explain products/services for customers.	Human rights & labor-related dimension
SRP2	Take care of customer complaints	Product responsibility-related dimension
SRS1	Invest to develop local community welfares Society-related dimension	Society-related dimension

Wikepedia.org describes the society-related ESG with human rights (similar to SRH1) and consumer protection (similar to SRP1). Investopedia.com describes the society-related ESG with customer care (similar to SRP2), employee health (similar to SRH3), and community engagement (similar to SRS1).

***Independent Variable: Governance-related ESG* (X_3)**

Stazyk et al. (2021) used age as a variable to measure different corporate governance policies on job satisfaction. The prosocial motivation of employees is directly impacted by the implementation of governance-related ESG. Data was collected for this survey from the Federal Employee Viewpoint Survey which confidentially asks questions about work experience, organization, and leadership using age as a key variable.

Code	Item	Reference
CG1	Considering everything, how satisfied are you with your Job?	Stazyk et al. (2021)
CG2	Considering everything, how satisfied are you with your Pay?	
CG3	Considering everything, how satisfied are you with your Organization?	

Kowske, et al. (2010) used age as a variable to measure the viewpoint of employee job satisfaction and work motivation in relation to skill recognition, job security, and career advancement. The study measured the variable through a WorkTrends employee opinion survey that has been run frequently every year since 1985.

Code	Item	Reference
CG3	Rate satisfaction with recognition you get for work	Kowske, et al. (2010)
CG4	Work gives a feeling of personal accomplishment	
CG5	Confidence in the future of the company	

To measure governance related ESG, we borrowed items from Woo (2013) as they address employee relations and diversity, corporate integrity through corruption avoidance, and financial transparency.

Code	Item	Reference
	I think the company I work for tries to:	
SRH2	Allow the freedom of labor union and forbid discrimination	Woo (2013)
SRS2	Avoid corruptions in business	
SREC1	Provide the company financial information to the public	

Control Variable: Prosocial Motivation - Gen Z and Gen Y Definition

Dhopade (2016) defines Gen Z as employees who were born between 1993 to 2011. Other studies describe Gen Z as employees born after 1994 (Bateh, 2019), born between 1997 to 2012 (Kasasa, 2021), born between 1996 to 2010 (Brown et al., 2019), or born between 1997 to 2013 (Schroth, 2019). Collectively, most of the literature suggests that Gen Z employees were born anytime between 1993 to 1997. It appears fair to use a middle value from 1993 to 1997 for the beginning of the Gen Z employee birth year. Accordingly, this research uses the middle value, 1995. Thus, employees are categorized into two generations for this study, as of April 2020, to be distinguished as follows:

- Gen Z: 18 to 24 years old
- Gen Y: 25 to 39 years old

Table 1. Variable & Items for Employee Prosocial Motivation

Variable	Code	Item	References
Prosocial Motivation (Y_1)	**PRO1**	Because I want to help others through my work	Grant and Berry (2011)
	PRO2	Because I want to have a positive impact on others	
	PRO3	Because it is important to me to do good for others through my work	
Environment-Related ESG (X_1)	**SREN1**	Take care of water, energy, and material uses	Woo (2013)
	SREN2	Minimize pollution when producing products/services	
	SREN3	Invest to protect the Environment	
Society-Related ESG (X_2)	**SRS1**	Invest to develop local community welfares	Woo (2013)
	SRS2	Avoid corruption in business	
	SRH1	Protect human rights at workplaces	
	SRH3	Clarify health care benefits for employees	

Governance-Related ESG (X_3)	**SRH2**	Allow the freedom of labor union and forbid discrimination	Woo (2013)
	SRS2	Avoid corruption in business	
	SREC1	Provide the company financial information to public.	

Regression Model

This study proposed to build an employee prosocial motivation model using regression analysis. Prosocial motivation was used for the dependent variable in the proposed regression model while the three ESG factors such as environment, society, and corporate governance served as independent variables. The multiple regression model was expressed as follows:

$$Y_1 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

where Y_1 = Prosocial Motivation
X_1 = Environmental ESG
X_2 = Social ESG
X_3 = Governance ESG

Sample Data

This research created a survey questionnaire using the variable items, based on the literature as shown in the summary table. The questionnaire was posted on Google Form at https://docs.google.com/forms/d/e/1FAIpQLSddB8Z8vVAmGP907gsrY8_EtaLYPqOXWPFigb2UZXQftr3BA/viewform. Survey data was collected using Amazon Mechanical Turk, requiring three conditions: that the Amazon Mechanical Turk workers had to be employed, aged from 18 to 55 years old, equal sampling among the three generations – Gen Z, Gen Y, and Gen X. The survey ran for a week in the third week of April 2020. The online survey received 220 responses. In the fourth week of August 2020, the survey was conducted again. The second survey collected 350 responses. Additional surveys were collected from senior students and MBA students in a public university in the New England region. Responses with multiple missing values or poor response quality were excluded from the reported data. There were 845 valid responses for this research.

4. RESULTS

Sample Description

The sample data includes a total of 845 responses, including 98 responses from Gen Z employees (11.6%), 386 responses from Gen Y (45.7%), 345 responses from Gen X employees (40.8%), and 16 responses from the Baby Boomers generation (1.9%). Refer to Table 4-1 below for the results of the sample data.

Table 4-1		
Generations		
	Frequency	**Percent**
Gen X (40-45)	345	40.8
Gen Y (25-39)	386	45.7
Gen Z (18-24)	98	11.6
Boomers (>=56)	16	1.9
Total	845	100.0

Descriptive Statistics & Correlations for All Data (N=845)

Descriptive statistical analysis shows that society (social)-related ESG has the highest mean score (5.377±1.148), followed by corporate governance (5.122±1.300) and environment-related ESG (4.486±1.565). The Pearson correlation analysis results show that employee prosocial motivation is highly correlated with the environment (.388), society (.387), and corporate governance related (.353) ESG variables ($p < 0.001$). Table 4-2 shows the results.

Table 4-2							
Descriptive Statistics and Correlations for All Data							
	Mean	**SD**	**N**	**(1)**	**(2)**	**(3)**	**(4)**
(1) Pro_Avg	4.933	1.667	845	1			
(2) X1_Environment	4.486	1.565	845	.388***	1		
(3) X2_Society	5.377	1.148	845	.387***	.563***	1	
(4) X3_Governance	5.122	1.300	845	.353***	.546***	.750***	1
Note: SD = Standard Deviation, *p < 0.05, **p < 0.01, ***p < 0.001 (2-tailed)							

Regression Results on All Data

This initial regression model tests all data points (N = 845) and includes three independent variables, environment, society (social), and corporate governance-related ESG, to estimate the dependent variable, employee prosocial motivation. In review of the model coefficients, it is shown that the model developed, [$R^2 = .194$, $R^2_{adj} = .191$, $F(3, 841) = 67.634$, $p < 0.001$] is statistically significant. The regression model explains 19.4% of the variance in the employee prosocial motivation outcome ($R^2 = 0.194$). Environment-related ESG and social-related ESG are significantly related to employee prosocial motivation ($p < 0.001$). Corporate

governance-related ESG exhibits no statistical significance. According to the standardized regression coefficient BETA, social-related ESG has the highest impact on employee prosocial motivation (.261), followed by environment-related ESG (.285). Table 4-3 reports the results on the full regression model for all generation employees.

Table 4-3

Regression Model Results for All Data

DV = Prosocial Motivation; R^2 = .194, R^2_{adj} = .191, F (3,841) = 67.634, p < 0.001; CI = 17.972; N = 845

	B	SE	BETA	*t-statistic*	*p-value*	VIF
(Constant)	1.768	.249		7.089	<.001	
X1_Environment	.250	.041	.234	6.099	<.001	1.543
X2_Society	.285	.071	.196	4.034	<.001	2.473
X3_Governance	.100	.062	.078	1.623	.105	2.408

Note: DV = Dependent variable; CI = Condition Index, B = Regression Coefficient, SE = Standard error, BETA = Standardized regression coefficient, VIF = Variance Inflation Factor

Descriptive Statistics & Correlations for Gen Z (N=345)

Descriptive statistical analysis shows that society (social)-related ESG has the highest mean score (5.377±1.148), followed by corporate governance (5.122±1.300) and environment related ESG (4.486±1.565). The Pearson correlation analysis results show that employee prosocial motivation is highly correlated with environment (.357), society (.416), and corporate governance related (.377) ESG variables (p < 0.001). Table 4-4 shows the results

Table 4-4

Descriptive Statistics and Correlations for Gen Z

	Mean	SD	N	(1)	(2)	(3)	(4)
(1) Pro_Avg	4.933	1.667	345	1			
(2) X1_Environment	4.486	1.565	345	.357***	1		
(3) X2_Society	5.377	1.148	345	.416**	.522**	1	
(4) X3_Governance	5.122	1.300	345	.377**	.530**	.772**	1

Note: SD = Standard Deviation, *p < 0.05, **p < 0.01, ***p < 0.001 (2-tailed)

Regression Results on Gen Z

The test of the second regression model looks specifically at Generation Z respondents (N=345). It includes all three independent variables (Environmental, Social, and Governance) of ESG and shows how it correlates with prosocial motivation. The Gen Z model is statistically significant, showed by the following: [R^2 = .202, R^2_{adj} = .195, F(3,341) = 28.815, p<.001]. The regression model explains 20.2% of the variance in prosocial motivation (R^2 = .202). Environment and Social related ESG are significantly related to prosocial motivation as their significance was .003 and .001, respectively. According to the standardized regression coefficient BETA, social-related ESG has the highest impact on employee prosocial motivation (.258), followed by environment-related ESG (.177). Governance-related ESG does not have a significant impact on prosocial motivation of Gen Z. According to Vittinghoff et al. (2012) and Kennedy (2003), there were no serious multicollinearity in the regression model as all VIFs were less than 10 and the condition index was less than 30.

Table 4-5						
Regression Model Results for Gen Z						
DV = Prosocial Motivation; R^2 = .202, R^2_{adj} = .195, *F (3*, 341) = 28.815, *p* <.001; CI = 17.392; N = 345						
	B	**SE**	**BETA**	***t-statistic***	***p-value***	**VIF**
(Constant)	1.549	.370		4.191	<.001	
X1_Environment	.207	.068	.177	3.040	.003	1.455
X2_Society	.359	.108	.258	3.319	.001	2.588
X3_Governance	.105	.098	.084	1.071	.285	2.620
Note: DV = Dependent variable; CI = Condition Index, B = Regression Coefficient, SE = Standard error, BETA = Standardized regression coefficient, VIF = Variance Inflation Factor						

Descriptive Statistics and Correlations on Gen Y Employees

The statistics show that society-related ESG has the highest mean score followed by governance-related ESG. Environment-related ESG has the lowest mean score. The scores are 5.434±1.107, 5.212±1.212, and 4.655±1.616, respectively. The Pearson correlation analysis shows that prosocial motivation is significantly correlated with Environmental (.386), Social (.342), and corporate governance (.299) ESG variables ($p < 0.001$).

Table 4-6							
Descriptive Statistics and Correlations for Gen Y							
	Mean	**SD**	**N**	**(1)**	**(2)**	**(3)**	**(4)**
(1) Mo_Int	4.945	1.635	386	1			
(2) X1_Environment	4.655	1.616	386	.386***	1		
(3) X2_Society	5.434	1.107	386	.342**	.582**	1	
(4) X3_Governance	5.212	1.212	386	.299**	.546**	.727**	1
Note: SD = Standard Deviation, *p < 0.05, **p < 0.01, ***p < 0.001 (2-tailed)							

Regression Results on Gen Y

The test of the third regression model looks specifically at Generation Y respondents (N=386). It includes all three independent variables (Environmental, Social, and Governance) of ESG and shows how it correlates with prosocial motivation. The Gen Y model is statistically significant, shown by the following: [R^2 = .170, R^2_{adj} = .163, F (3,382) = 26.075, p<.001]. The regression model explains 17% of the variance in prosocial motivation (R^2 = .170). Prosocial motivation is significantly related to Environmental ESG ($p < 0.001$) and Social ESG ($p < 0.05$). Governance-related ESG, shows no statistical significance (p=.616). Governance-related ESG does not have a significant impact on the prosocial motivation of Gen Y. According to the standardized regression coefficient BETA, environmental ESG has the highest impact on employee prosocial motivation (.276), followed by social (.156). According to Vittinghoff et al. (2012) and Kennedy (2003,) there was no serious multicollinearity in the regression model as all VIFs were less than 10 and the condition index was less than 30.

Table 4-7						
Regression Model Results for Gen Y						
DV = Prosocial Motivation; $R^2 = 0.170$, $R^2_{adj} = 0.163$, F (3 ,382) = 26.075, p = <.001; CI = 18.29; N = 386						
	B	**SE**	**BETA**	**T-statistic**	**p-value**	**VIF**
(Constant)	2.149	.391		5.497	< .001	
X1_Enviornment	.279	.059	.276	4.695	< .001	1.589
X2_Society	.230	.106	.156	2.173	.030	2.365
X3_Governance	.047	.094	.035	.501	.616	2.227
Note: DV=Dependent Variable; CI = Condition Index, B = Regression Coefficient, SE = Standard Error; BETA = Standardized Regression Coefficient, VIF = Variance Inflation factor						

Summary Table of Regression Results

Table 4-8			
Summary of Regression Models			
Model	**Full (All Data)**	**Gen Z**	**Gen Y**
DV	Prosocial Motivation	Prosocial Motivation	Prosocial Motivation
Environment	.234***	.177**	.276***
Society	.196***	.258**	.156*
Governance	.078	.084	0.035
R2	.194	.202	.170
R2adj	.191	.195	.163
F	67.634	28.815	26.075
N	845	345	386
Note: #p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001; Standardized regression coefficient			

5. DISCUSSION

Hypothesis 1a

Results of full data analysis supported Hypothesis 1a, such that Environment-related ESG is positively related to Employee Prosocial Motivation. The results show a BETA of .234 with a p-value of ($p<0.001$) indicating a strong significance.

These results are in line with findings presented by Sipahutar et al. (2022) which studied the effects of workers' motivation when working to improve environmental problems and observed an increase in work motivation can positively affect environmental issues surrounding specific types of work. These results are further supported by a study conducted by Amy Yeo Chu May et al. (2021) which performance analysis to determine whether an organization's environmentally sustainable activities when communicated to employees within the organization motivate employees. The study found that employees who work for a company that is environmentally sustainable have a positive effect on employees' level of organizational trust. Additional prior studies are consistent with the effects of workers' motivation (Chein et al., 2020 & Suganthi et al., 2019).

Prosocial motivation is the idea that an individual is motivated by performing activities or providing services or products which benefit the greater whole. Employees exhibit a higher level of prosocial motivation when working for companies that are environmentally responsible and have sustainable practices in place. If an employee feels the work, they are doing is beneficial for the environment and that others are benefiting from the work activity, they will be highly motivated to continue doing such work.

Hypothesis 1b

Evidence does not support Hypothesis 1b in that generations do not moderate the effect of environmental-related ESG on employees' prosocial motivation ESG; environmental-related ESG was not more related to the prosocial motivation of younger generations. Gen Y employees (BETA =.276, $p < 0.001$) are more impacted by environment-related ESG on their prosocial motivation in the workplace than Gen Z employees (BETA=.177, $p < 0.01$).

The results are consistent with the research conducted by Chatzopoulou and Kiewiet (2020), who studied how younger generations evaluated corporate social responsibility. While in the study millennials did have a deeper understanding of ethical complications like climate change, a large proportion of respondents did not identify as ethical, which inherently contributed to a lack of prosocial motivation. Hassan and Kodwani (2020) Also found similar results, stating that younger generations are apprehensive about trusting the environmental CSR claims of companies in the first place and that they typically do not affect motivations to work for a company with such claims.

The Chatzopoulou and Kiewiet (2020) article noted that there is a broad cynicism about ethicality that may cause a reversed effect on social desirability; respondents may not want to be associated with an attitude that they view as not authentic, and this case may view the climate implications of an organization's actions as necessary for business. Authenticity is a major factor in motivating younger generations; many younger workers view CSR implementations as potentially inauthentic and are only implemented for the sole purpose of boosting public image.

Hypothesis 2a

From the full model across all surveyed generations (N=845), society (social)-related ESG presents with a <.001 p-value across both the correlation and multiple regression model results, including a .196 standardized regression coefficient (BETA) value from the regression results and a .387 correlation value. Therefore, evidence supports hypothesis H2a wherein social-related ESG is positively related to prosocial motivation. Further, from this evidence, the results of the study show that social-related ESG is statistically significant as related to prosocial motivation (p<.001).

The results regarding the connection between social-related ESG and prosocial motivation are consistent with prior studies (Kajackaite et al., 2020; Bernardino et al., 2021; Fraizer & Tupper et al., 2018; Skudiene & Auruskevicien et al., 2010). Prior research suggests that companies that are deeply connected to their corporate social responsibility efforts tend to have better prosocial motivation. Similarly, employee thriving had a positive relationship with prosocial motivation. Collectively, these results in combination with the outcomes of prior studies suggest that various positive attributes related to social-related ESG can be integrated into the work environment to positively influence prosocial motivation.

Social-related ESG focuses on the establishment and maintenance of relationships between a company and its stakeholders. Social relationships thrive in supporting environments that positively enhance and sustain the relationship in question. Further, maintaining positive social relationships creates opportunities for initiative and growthforo those connected to the relationship. However, unsuitable environments that negatively establish and maintain the relationship in question will inevitably wither away. Businesses that are socially responsible, create socially responsible programs and initiatives to enhance the social-related ESG of the external and internal community will strengthen employees' efforts, engagement, satisfaction, and prosocial behaviors.

Hypothesis 2b

Evidence does support Hypothesis 2b in that society-related ESG has a positive impact on employees' prosocial motivation, with the impact being more positively related to the prosocial motivation of younger generations. Gen Z employees (.258, $p < 0.001$) get more impacts of society-related ESG on their prosocial motivation in the workplace than Gen Y employees (.156, $p < 0.05$).

These results are consistent with the research conducted by Rank and Contreras (2021) which found a correlation between employee age and workplace motivation when focusing on CSR initiatives within a company. It was noted that younger generations are impacted the most when considering a company's corporate social responsibility practices. Sinha et al. (2021) also found similar results, as it was noted that younger generations report a stronger desire to have practices of society-related ESG, such as social needs, stability, and physiological needs to improve workplace motivation. There was an emphasis on younger generations preferring these practices at a higher degree than older generations. Wisse et al. (2015) also found there to be generational differences in workplace motivation based on CSR.

In looking at why younger generations are prosocially motivated to work for companies that place a high value on society-related ESG, it is important to note the difference in the upbringing of these younger generations. Rapid societal change has occurred to the extent that schools, media, and other frequently accessed forms of communication have placed a strong emphasis on societal reformation where a newer form of thinking carries over into the foundations of what the younger generations want in their work. This inevitably leads to an increase in the

work motivation of younger generations deriving from organizations' impacts on their surrounding communities.

Hypothesis 3a

Evidence does not support Hypothesis 3a, governance-related ESG is positively correlated to prosocial motivation in the workplace. The results show no statistical significance between governance-related ESG and prosocial motivation (BETA=.078, p = 0.105).

The results are consistent with Moore (2004) who reported insignificant findings of a correlation between motivation and governance. Landry et al. (2017) reported that cash incentives are the best use for motivating employees. Both studies have shown that governance-related ESG and prosocial motivation do not fit together.

When comparing the three aspects of ESG; Environmental, Societal, and Governance, it seems easy to point out that governance would fall below the other two. With everything going on in the world, it is important to protect and preserve the environment. The next most important aspect is giving back to the communities and society. Every company wants to show they care about social issues that could affect their employees. When stacking those two aspects against governance, it makes sense that governance comes back as insignificant. If we were only doing a study based only on environment-related ESG, then our results may tell a different story. There are all different types of motivations and ways to motivate employees. This study shows that some weigh more heavily on employees than others do.

Hypothesis 3b

The evidence shows weak support for hypothesis 3B, governance-related ESG has a larger impact on prosocial motivation in Gen Z employees than in Gen Y Employees. Even though the data is statistically insignificant between both Gen Z and Gen Y, it is hard to ignore that Gen Z has a correlation more than 3 times higher than that of their Gen Y counterparts (BETA = .084 and .035, respectively).

Our results are consistent with the findings of Srivastava and Banerjee (2016) and Made, et al (2018). They both reported a strong correlation between governance-related ESG and work motivation in younger employees.

As an employee moves throughout their career, their goals and motivations change. Older generations are usually more conservative and like the status quo. They value a stable environment that does not have much change. They are not motivated by the whole company's structure. They are usually more focused on what affects them day to day than what could impact them year over year. Younger generations, however, want to come in and make an impact. If a young employee has goals to reach senior leadership, then the way the company is run would directly motivate them to work in a certain way. Their goals are tied to the structure and decisions of the company.

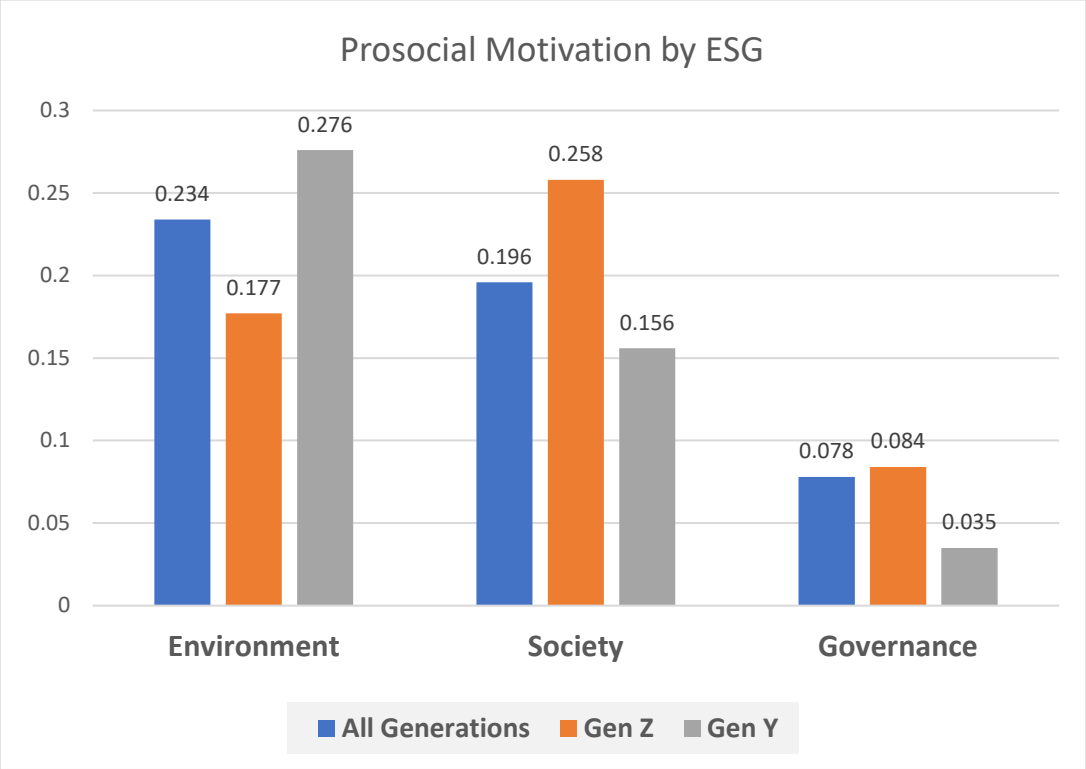


Table 5-1

Summary of Findings

Hyp	DV	IV	Moderator	Explanation
H1a	Prosocial Motivation	Environment ESG	None	**Supported**; BETA value of .234 with strong significance. Employees have high prosocial motivation with an environmentally conscious employer
H1b	Prosocial Motivation	Environment ESG	Generation	**Not Supported**: Data shows Gen Y employees (.276, $p < 0.001$) are more impacted by Environment-related ESG on their prosocial motivation in the workplace than Gen Z employees (.177, $p < 0.01$).
H2a	Prosocial Motivation	Society ESG	None	**Supported:** Beta value of .196 showing significance. Positive correlation between employee prosocial motivation and social-related ESG with statistically significant outcomes.

H2b	Prosocial Motivation	Society ESG	Generation	**Supported:** Beta value of .258 for Gen Z, and .156 for Gen Y. There is a positive correlation between younger generations and society-related ESG.
H3a	Prosocial Motivation	Governance ESG	None	**Inconclusive:** The results show no statistical significance between governance-related ESG and prosocial motivation (.078, $p<0.1$)
H3b	Prosocial Motivation	Governance ESG	Generation	**Partially Supported:** The evidence shows weak support for hypothesis 3b. Gen Z has a correlation more than 3 times higher than that of their Gen Y counterparts (.084 and .035, respectively
Note: Prosocial Motivation = Employee work / prosocial motivation				

Managerial Implications

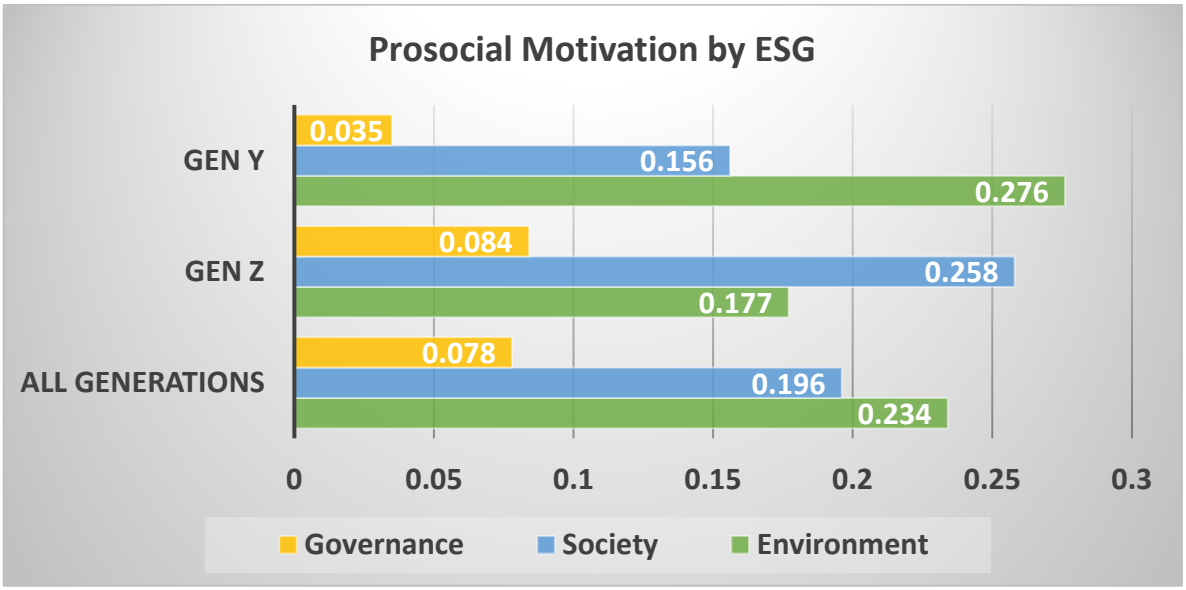
Overall, the results of this research will benefit businesses looking to strengthen their employees' prosocial motivators and improve business-specific factors that lead to increased profitability and employee engagement and satisfaction. The intent of this research is to highlight how the three ESG factors are strategically used to increase employee motivators with a specific focus on the level of influence each factor has on each generation (Gen Z and Gen Y). By incorporating this research into business practice, businesses will be able to implement environmental or societal-related programs by generations to increase productivity and prosocial motivation.

In general, the business organization should prioritize Environment-related ESG among the three factors. The next factor that would be prioritized would be Social-related ESG. The results showed that Governance-related ESG was not significant, so this should not be prioritized.

For Gen Z, the regression model shows the business should prioritize Society-related ESG and then Environment-related ESG. Governance-related ESG was insignificant as well for Gen Z and this should not be prioritized by the business. It is important to reference the switch between Gen Z and all generations. Society-related ESG is at the top of the list for Gen Z compared to Environment for all generations. This is especially important to note when trying to motivate younger generations.

For Gen Y, the regression model shows the business should prioritize Environment-related ESG followed by Society-related ESG in terms of having a positive impact on Generation Y Prosocial Motivation. Governance-related ESG was not statistically significant as it did not have a strong impact on prosocial motivation for generation Y workers. Given this information, it should not be prioritized by business leadership.

Table 5-2			
Executive Recommendation for Employee Prosocial Motivation with ESG			
Priority	**All Generations**	**Gen Z**	**Gen Y**
1	Environment	Society	Environment
2	Society	Environment	Society



6. CONCLUSION

This research examines the impact of environment-related ESG, social-related ESG, and corporate governance-related ESG on employee prosocial motivation. The multiple regression model is used to analyze the relationship between each ESG factor on prosocial motivators. For all generations, environment-related ESG is identified as having a statistically significant relationship to employee prosocial behavior, followed by social-related ESG. Gen Y employees are found to share the same ranking as those ESG factors. Gen Z, the youngest group of employees in a workplace, considers environmental-related ESG to be more impactful to prosocial motivation, followed by social-related ESG factors. Across all models, this research finds no relationship to support that corporate governance-related ESG has a significant impact on prosocial motivation.

Limitations of Current Research

There were multiple limitations to note while conducting this research study, most of which were data related. Due to time and funding constraints, the research can only explore employee prosocial motivation related to generations Y and Z as the moderator and three independent variables. While the sample data included a total of 845 responses, there were only 98 responses from Gen Z employees (11.6%) and 386 responses from Gen Y (45.7%). Given this large difference, this could have impacted the results of the study since Gen Z employees were not as well represented. To improve this research study, a larger more comprehensive population could be surveyed such that each generation has equal representation in the study. Increasing the time allowed to complete the survey and opening it to other geographic locations would help ensure results are not US-specific.

Closing Remarks

This research provides insight into workplace prosocial motivation by providing empirical evidence on the differences in the impact that the areas of ESG have. The evidence suggested that decisions companies make in promoting environment-related and social-related ESG into their culture have a considerable impact on the prosocial motivation of younger generations. This means that businesses should prioritize these two areas of ESG when prosaically motivating younger generations within their workforce.

REFERENCES

Abid, G., Sajjad, I., Natasha, S. E., Farooqi, S., & Nisar, A. (2018). The influence of prosocial motivation and civility on work engagement: The mediating role of thriving at work. *Cogent Business & Management, 5*(1), Article Number 1493712. DOI:10.1080/23311975.2018.1493712

Alonso-Almeida, MDM, Llach, J. (2019). Socially responsible companies: Are they the best workplace for millennials? a cross-national analysis.; *26(1),* 238– 247. DOI:10.1002/csr.1675

Amy Yeo Chu May, Gan Soong Hao, & Carter, S. (2021). Intertwining corporate social responsibility, employee green behavior, and environmental sustainability: the mediation effect of organizational trust and organizational identity. *Economics, Management & Financial Markets*, *16*(2), 32–61. https://www.proquest.com/docview/2548712998/fulltextPDF/838339802953415DPQ/1?accountid=9970

Armstrong-Stassen, M., & Schlosser, F. (2008). Benefits of a supportive development climate for older workers. *Journal of Managerial Psychology*, *23*(4), 419–437. DOI:10.1108/02683940810869033

Bernardino, P. (2021). Engaging employees through corporate social responsibility programs: Aligning corporate social responsibility and employee engagement. *Journal of Organizational Psychology, 21*(1), 105-113. DOI:2512307936/se-2

Bilge, H., Virlanuta, F. O., Zungun, D., Barbuta-Misu, N., Comuk, P., & Guven, E. T. A. (2021). Generation y's perception of servant leadership and job satisfaction. *Economies*, *9*(1), 24. DOI:10.3390/economies9010024

Calk, R., & Patrick, A. (2017). Millennials through the looking glass: workplace motivating factors. *Journal of Business Inquiry: Research, Education & Application*, *16*(2), 131–139. https://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=123219586&site=ehost-live&scope=site

Chan, H. L., Kawada, B., Shin, T., & Wang, J. (2020). CEO-employee pay gap and firm R&D efficiency. *Review of Accounting & Finance*, *19*(2), 271–287. DOI:10.1108/RAF-10-2018-0207

Chatzopoulou, E., & Kiewiet, A. (2021). Millennials' evaluation of corporate social responsibility: The wants and needs of the largest and most ethical generation. *Journal of Consumer Behaviour*, *20*(3), 521–534. DOI:10.1002/cb.1882

Chen, C.-F., Chen, S.-C., & Tsai, P.-S. (2020). Job passion in the context of green restaurant: Environmental stewardship orientation and job autonomy as antecedents. Journal of General Management, 46(1), 16–25. DOI:10.1177/0306307020911177

Chien, G. C. L., Mao, I., Nergui, E., & Chang, W. (2020). The effect of work motivation on employee performance: Empirical evidence from 4-star hotels in Mongolia. *Journal of Human Resources in Hospitality & Tourism*, *19*(4), 473–495. DOI:10.1080/15332845.2020.1763766

Day, J., Karani, A., Adler, H., & Nicely, A. (2013). corporate social responsibility and college recruiting in the hospitality industry. *Journal of Human Resources in Hospitality & Tourism*, *12*(1), 71–90. DOI:10.1080/15332845.2013.723269

Damayanti, R. D. A., Indrawati, R., & Aida, M. (2021). The mediating role from work motivation towards the implementation of corporate governance and rsud x's

performance. *Dinasti International Journal of Education Management And Social Science*, 3(1), 46-61. DOI:10.31933/dijemss.v3i1.987

Einolf, C. (2016). Millennials and public service motivation: findings from a survey of master's degree students. *Public Administration Quarterly,* 40(3), 429–457. http://www.jstor.org/stable/24772878

Fantini, L., Leiendecker, J., Gehra, B., Weisskopf, A., & Varshney, A. (2022). ESG compliance in an era of tighter regulation. *Boston Consulting Group.* DOI:2715894864/se-2

Frazier, M. L., & Tupper, C. (2018). Supervisor prosocial motivation, employee thriving, and helping behavior: a trickle-down model of psychological safety. *Group & Organization Management,* 43(4), 561–593. DOI:10.1177/1059601116653911

Garrett, J. B., Holderness Jr., D. K., & Olsen, K. J. (2021). An experimental investigation of how self-interested organizational norms undermine prosocial motivation and influence employee effort. *Journal of Information Systems*, *35*(2), 17–36. DOI:10.2308/ISYS-19-047

Glavas, A., & Piderit, S. K. (2009). How does doing good matter?: Effects of corporate citizenship on employees. *Journal of Corporate Citizenship*, 36(1), 51–70. DOI:10.9774/GLEAF.4700.2009.wi.00007

Global Reporting Initiative (GRI). (2012). Global Reporting Initiative Sustainability Report 2010/11. Retrieved June 5, 2020 from https://www.globalreporting.org/resourcelibrary/GRI-Sustainability-Report-2010-2011.pdf

Grant, A. M. & Berry, J. W. (2011). The necessity of others is the mother of invention: Intrinsic and prosocial motivations, perspective taking, and creativity. Academy of Management Journal, 54(1), 73-96. DOI:10.5465/amj.2011.59215085

Guerrero, M., Amorós, J. E., & Urbano, D. (2021). Do employees' generational cohorts influence corporate venturing? A multilevel analysis. *Small Business Economics*, *57*(1), 47–74. DOI:10.1007/s11187-019-00304-z

Hassan, Y., & Kodwani, A. D. (2020). Assessing the influence of spirituality and religiosity in choosing a socially responsible employer: the case of Generation Z. South Asian Journal of Management, 27(3), 79-110. https://search-ebscohost.org/login.aspx?direct=true&db=buh&AN=147916841&site=ehost-live&scope=site

Hur, W.-M., Moon, T.-W., & Ko, S.-H. (2018). How employees' perceptions of CSR increase employee creativity: mediating mechanisms of compassion at work and intrinsic motivation. *Journal of Business Ethics*, *153*(3)*,* 629–644. DOI:10.1007/s10551-016-3321-5

The Investopedia Team. (2022). *Environmental, social, and governance (ESG) criteria*. Investopedia. https://www.investopedia.com/terms/e/environmental-social-and governance-esg-criteria.asp

Klimkiewicz, K., & Oltra, V. (2017). Does CSR enhance employer attractiveness? The role of millennial job seekers' attitudes. *Wiley Online Library*, *24*(5), 449-463. DOI:10.1002/csr.1419

Kunz, J. (2020). Corporate social responsibility and employees motivation—broadening the perspective. *Schmalenbach Business Review*, *72*(2), 159–191. DOI:10.1007/s41464-020-00089-9

Kowske, B., Rasch, R., & Wiley, J. (2010). Millennials' (lack of) attitude problem: an empirical examination of generational effects on work attitudes. *Journal of Business & Psychology*, *25*(2), 265–279. DOI:10.1007/s10869-010-9171-8

Landry, A. T., Forest, J., Zigarmi, D., Houson, D., & Boucher, E. (2017). The carrot or the stick? investigating the functional meaning of cash rewards and their motivational power according to self-determination theory. *Compensation & Benefits Review*, 49(1), 9–25. DOI:10.1177/0886368717750904

Lanfranchi, J., & Narcy, M. (2022). How do prosocial motivation and performance-related pay interact in the workplace context? Evidence from the non-profit sector. *Kyklos, 75*(3), 436–455. DOI:10.1111/kykl.12297

Liao, H., Su, R., Ptashnik, T., & Nielsen, J. (2022). Feeling good, doing good, and getting ahead: A meta-analytic investigation of the outcomes of prosocial motivation at work. *Psychological Bulletin, 148*(3), 158-198. DOI:10.1037/bul0000362

Luger, M., Hofer, K. M., & Floh, A. (2022). Support for corporate social responsibility among generation Y consumers in advanced versus emerging markets. *International Business Review*, *31*(2), N.PAG. DOI: 10.1016/j.ibusrev.2021.101903

Mahmoud, Fuxman, L., Mohr, I., Reisel, W. D., & Grigoriou, N. (2021). "We aren't your reincarnation!" workplace motivation across X, Y and Z generations. *International Journal of Manpower, 42(*1), 193–209. DOI:10.1108/IJM-09-2019-0448

Moore, J. (2004). On scurvy,elephants, and governance vitamins. *Board Leadership*, *2004*(72), 4–6. DOI:10.1002/bl.38620047203

Ramalaksana Made, Sukmawati Anggraini, & Hubeis Aida Vitayala S. (2018). The influence of personal value and work motivation on employee engagement in Generation y. *Russian Journal of Agricultural and Socio-Economic Sciences*, è78(6), 346-354. DOI:10.18551/rjoas.2018-06.40

Rank, S., & Contreras, F. (2021). Do millennials pay attention to corporate social responsibility in comparison to previous generations? Are they motivated to lead in times of transformation? A qualitative review of generations, CSR and work motivation. *International Journal of Corporate Social Responsibility*, *6*(4).1-13 DOI:10.1186/s40991-020-00058-y

Shin, I, Hur, W-M. (2020) How are service employees' perceptions of corporate social responsibility related to their performance? Prosocial motivation and emotional labor as underlying mechanisms. *Corporate Social Responsibility and Environmental Management*. 27(6): 2867– 2878. DOI: 10.1002/csr.2008

Shin, Y., & Hur, W. (2021). When do job-insecure employees keep performing well? the buffering roles of help and prosocial motivation in the relationship between job insecurity, work engagement, and job performance. *Journal of Business and Psychology, 36*(4), 659-678. DOI:10.1007/s10869-020-09694-4

Sinha, S., & Vispute, S. (2022). Workplace motivation and job satisfaction among generation Y in india. *Cardiometry, 22(1),* 251-257. DOI:10.18137/cardiometry.2022.22.251257

Shakil, H. (2020). The effect of leadership and motivation on employees satisfaction: evidence from mirpur division aj & k in restaurant sector. *International Journal of Management, Accounting & Economics*, *7*(8), 435–447. https://www.ijmae.com/article_114307_353277c491c7a5519ced04f8c46b31fb.pdf

Sipahutar, Y. H., Rahmayanti, H., Achmad, R., & Sitorus, R. (2022). increased effectiveness of conservation the coastal environment through cleaner production and work motivation of

fish processors. *IOP Conference Series*. *Earth and Environmental Science*, *950(*1), 012050. DOI:10.1088/1755-1315/950/1/012050

Skudiene, V., & Auruskeviciene, V. (2012). The contribution of corporate social responsibility to internal employee motivation. *Baltic Journal of Management, 7*(1), 49-67. DOI:10.1108/17465261211197421

Solnet, D., Kralj, A., & Kandampully, J. (2012). Generation Y employees: An examination of work attitude differences. *Journal of Applied Management and Entrepreneurship, 17*(3), 36-54. Retrieved from https://idm.oclc.org/login?url=https://www.proquest.com/scholarly-journals/generation-y-employees-examination-work-attitude/docview/1041243635/se-2

Srivastava, M., & Banerjee, P. (2016). Understanding Gen Y. *Journal of Management Research*, *16(*3), 148–164.

Suganthi, L. (2019). Examining the relationship between corporate social responsibility, performance, employees' pro-environmental behavior at work with green practices as mediator. *Journal of Cleaner Production*, 232(1), 739–750. DOI: 10.1016/j.jclepro.2019.05.295

Stazyk, E. C., Davis, R. S., & Liang, J. (2021). Probing the links between workforce diversity, goal clarity, and employee job satisfaction in public sector organizations. *Administrative Sciences, 11(3),* 77. DOI:10.3390/admsci11030077

Wisse, et al. (2018) catering to the needs of an aging workforce: The role of employee age in the relationship between corporate social responsibility and employee satisfaction. *Journal of Business Ethics,* 147(4), 875–888. DOI:10.1007/s10551-015-2983-8

Wong, M., Gardiner, E., Lang, W., & Coulon, L. (2008). Generational differences in personality and motivation Do they exist and what are the implications for the workplace? *Journal of Managerial Psychology*, *23*(8), 878–890. DOI:10.1108/02683940810904376

Wong, W. C., Ahmad, A. H., Mohamed-Arshad, S., Nordin, S., & Adzis, A. A. (2022). Environmental, social and governance performance: Continuous improvement matters. *Malaysian Journal of Economic Studies, 59*(1), 49-69. DOI: 10.22452/MJES.vol59no1.3

Woo, H. (2013). Do consumers want a "good" apparel brand? The effects of apparel brands' corporate social responsibility (CSR) practices on brand equity moderated by culture. Unpublished master's thesis, University of North Carolina at Greensboro.

Woo, H., & Jin, B. (2012). Corporate social responsibility communications of global apparel firms: The website analysis of U.S., European, and Asian firms. In Ahn, C., & Burns, L. D. (Co-chairs), *KSCT-ITAA Joint Symposium: Enhancing International Collaborations.* Symposium conducted at the annual conference of International Textile and Apparel Association (ITAA), Hawaii, United States.

Yang, J., & Basile, K. (2022). Communicating corporate social responsibility: External stakeholder involvement, productivity and firm performance. *Journal of Business Ethics*, *178*(2), 501-517. DOI:10.1007/s10551-021-04812-5

Yu, C., Ye, B., & Ma, S. (2021). Creating for others: linking prosocial motivation and social entrepreneurship intentions. *Management Decision*, *59*(11), 2755–2773. https://www.emerald.com/insight/0025-1747.htm

Van der Voet, J., & Steijn, B. (2021). Relational job characteristics and prosocial motivation: A longitudinal study of youth care professionals. *Review of Public Personnel Administration*, *41*(1), 57–77. DOI:10.1177/0734371X19862852

Zainee, Intan Azurin, and Fadilah Puteh (2020). Corporate social responsibility impact on talent retention among generation Y. *Revista De Gestão*, *27*(4), 369–392., DOI:10.1108/rege-06-2019-0070.

Modeling Fast Fashion: Sustainability of Complementary Products

ABSTRACT

Investigations in the sustainability of the fast fashion industry have been heavily researched, specifically in the clothing manufactured and the environmentally harmful materials used in production. This study addresses the complementary products sold alongside fast fashion garments. The analysis focuses on the equally significant impacts of accessories and cosmetics have in a rapidly shifting industry environment due to cultural shifts that have shortened the life span of trend cycles. These cultural influences have rendered fast fashion products time-sensitive and prone to low production quality and constant disposal. Utilizing circular economy methods and increasing scalability of these sustainable solutions may lessen the hidden effects complementary products may have on the environment, since their operations and production are typically ignored under the main retail brand.

Keywords: Complementary Products, IPAT Equation, Circular Economy, Time-Sensitive, Social Trends.

INTRODUCTION

In recent years, the retail fashion industry, specifically concerning fast fashion and major international brands such as SHEIN, H&M, and Zara, has seen a spike in demand and production in clothing and accessories. In opposition to a traditional fashion model, these companies' operations are influenced by dynamic trend cycles. Traditional fashion models are characterized by higher production costs and wages, higher quality and ethically sourced materials, longer lasting, more timeless designs and are usually manufactured on a more local scale.

Figure 1. Drivers of increased environmental pollution in the fast fashion industry

Figure 1 provides insight to drivers of increased consumerism and fast fashion turnover through an investigation on the consumer and industry environment of fast fashion, heavily contrasting the traditional model. Efforts to satisfy expanding customer demands due to easier customer reachability which increases the products' perceived value and ultimately, increases in disposable income allows customers to purchase more fast fashion garments as the trend sees fit. In turn, these higher demands and adherence to trend cycles contribute to high levels of inventory turnover and render fast fashion products extremely time sensitive.

Another factor to consider in the sustainability of fast fashion products is the environment in which the product is created. A cultural shift to rapidly overturned trends has also shifted the operation of retail supply chains to accommodate these trend cycles.

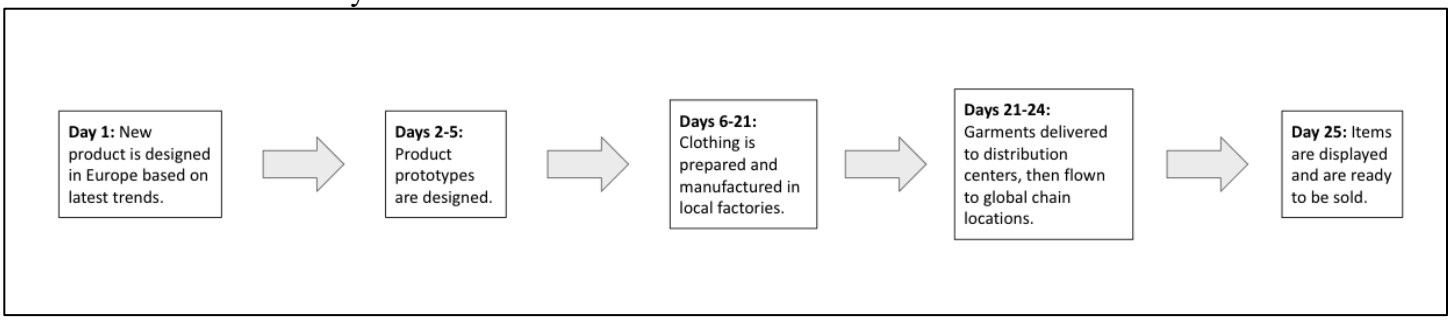


Figure 2. Zara fast fashion product supply chain [15]

Figure 2 exemplifies the ultra-fast pace that Zara, a leading brand in the fast fashion industry, follows in producing new collections to match the speed of changing trend cycles. New products can be sourced from the current trends and appear on fast fashion shelves within 25 days. The constant arrival of fresh garments assumes the obsolescence and disposal of the previous collection of clothing, proving its time sensitivity, and produces higher quantities of waste opposed to traditional seasonal collections. Figure 1 displays the consequences of this fast fashion model; products are created with lower quality and lack of local production, sustainable material used, and lower worker compensation.

The circular economy model is an alternative option for more sustainable practices in the rapidly growing fast fashion realm. The model focuses on minimalizing waste by creating products with reusability in mind once discarded [16]. Fast fashion garments are typically rendered obsolete and are trashed once the trend has ended, and garments created under the circular economy model will permit less waste if the irrelevant products can be regenerated into the next generation of clothing.

The study addresses the research gap in the complementary products sold alongside the main designer garments manufactured in major retail companies such as SHEIN, H&M, and Zara amongst other smaller retailers in the industry. The related literature offers a wide range of studies concerning the impact fast fashion clothing items have on sustainability and the drivers behind increasing consumerism. However, only a limited number of these acknowledge the influence of complementary products which is just as significant. Aiming at providing a viable model, this study focuses on complementary products such as bags, accessories, and cosmetics, that accompany the main garments sold under the same overarching brand name.

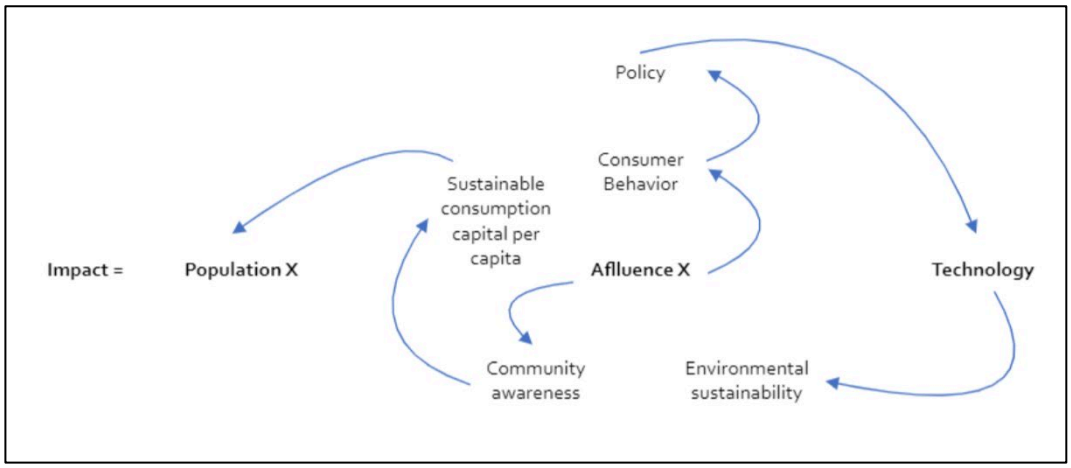


Figure 3. Model detailing environmental impact of fast fashion, based on IPAT equation

Figure 3 uses the IPAT equation (Environmental Impact = Population x Affluence x Technology) as a baseline while considering other factors specific to the growing consumerism and impact from the fast fashion industry that contribute to each variable, also taking industry and consumer environment into consideration.

LITERATURE REVIEW AND BACKGROUND

The global fashion industry produces close to two billion tons of carbon emissions, between 3 to 5 percent of the global amount, an environmental impact that accounts for more than the aviation and shipping industries combined [1]. The impact of the industry is more substantial in its fast-fashion segment where the efforts are concentrated on time-sensitive, highly responsive products with shortened time to market.

Fast-fashion products meet the customer demand for high-end or luxury-fashion products at significantly lower costs [2] while making use of economies of scale via mass standardization or customization. Often associated with relatively low-quality levels, fast-fashion products have shorter life expectancies as a result of conscious pricing strategies [3] and high obsolescence rates due to their time-sensitive nature. Coupled with the increasingly rapid changes in consumer preferences and social trends these characteristics prepare the perfect conditions for their rapid disposal.

Despite the efforts in creating sustainable solutions for their sourcing, fabrication, reuse, and recycling, etc., fashion products are customarily considered to be unsustainable. SHEIN, H&M, Zara, are three major retailers with their collective almost 60 percent market share. These international fast-fashion clothing retailers offer clothing along with bags, accessories, and cosmetics to consumers. The demand for their products is heavily reliant on the informational and cultural content creation [4], and often fueled by increasing consumerism. These organizations and other similar businesses are frequently scrutinized by the public for encouraging over consumption and their adverse impact on the environment. Although they are considered to be laggards [5] in their adaptation of sustainable practices, some are successfully able to utilize sustainability as a competitive advantage given that majority of the consumers consider environmental impact as a significant factor in their purchase decisions [6,7].

Advancements in technology and social media have allowed these companies to more easily market and sell their products and as a result, demand according to new trends have increased. Ultimately, garments and products must move through supply chains at a much faster rate to produce to meet these demands dictated by

social trends. To ensure sustainable conditions within these supply chains, organizations must aim to monitor smaller producers more closely down their supply chain and encourage compliance with company sustainability rules of conduct [12].

Renewable crops, organic materials, and carbon positive are some of the sustainability practices that relate to transparency, sourcing policies, governance [8], and producers' responsible production plans in fast fashion industry. In addition to sourcing and other operations, new business plans that focus on consumer types, behaviors, and actions [9] also emerge with a goal of extending the life cycle of products via slow fashion [10], remake, repair, rent, reuse, recycle and other environmentally friendly options. Additional attempts focus on purchasing decisions and consumer behavior in fashion industry, recommend reshaping the trends of fashion towards sustainability.

In addition to company's commitment to reducing greenhouse emission, achieving carbon neutrality and other climate targets, increasing scalability of innovative solutions for sustainable and cost-efficient products is one effective way of ensuring sustainability of fast fashion industry. Aiming to follow a circular economy model may offer yet another solution for long-term sustainability in what otherwise is an industry built on short life expectancy and brisk product turnover. Designed for material recovery and reuse [13], shifting towards a circular model where rapidly obsolete products are prematurely disposed of after trends have passed will ensure minimal waste.

Despite the move towards circularity in retail industry and the abundancy of studies investigating sustainability of fast-fashion retailers, only a few of them concentrate on the impact of complementary products on the environment and society. That is, majority of the studies concentrates on clothing purchases and garment disposal, while the sales of bags, accessories, and cosmetics of these global companies, however, can reach up to $300 bn [11] with an expected growth over 11 percent annually.

For instance, jewelry pieces manufactured under a fast fashion model are created to follow trends similar to clothing. With trend cycles accelerated to up to 12 collections per year versus a traditional fashion brand's seasonal cycle, many companies have imitated this rapid adaptation to fleeting consumerism in its jewelry and accessories collection. German jewelry company "Beeline" satisfies these trends by adding hundreds of new accessories to their collection per year, opposed to other fine-jewelry companies producing an average of merely two collections per year [14].

This study focuses on what will be referred as "complementary" fast-fashion merchandise which consists of bags, accessories, cosmetics that carry the main brand name.

CONCLUSION

Due to increases in the rate of social trends and marketing and operational efficiency, major retail brands have shifted to manufacturing for mass consumerism and as a result, waste generation has increased, and sustainability has become challenging to manage in this environment. Further investigations on the effects of complementary products, in addition to main brand garments, is necessary to fully attend to sustainable solutions with such quick product turnover and relevance. Increasing scalability of sustainable practices and shifting to a circular economy are only some of the solutions encouraged to fully.

REFERENCES

1. Achim Berg, Anna Granskog, Libbi Lee, and Karl-Hendrik Magnus, Fashion on climate, Report, McKinsey, August 26, 2020, Accessed: August 16, 2022.
2. Barbara Cimatti, Giampaolo Campana, Laura Carluccio, Eco Design and Sustainable Manufacturing in Fashion: A Case Study in the Luxury Personal Accessories Industry, Procedia Manufacturing, Volume 8, 2017, Pages 393-400, ISSN 2351-9789, https://doi.org/10.1016/j.promfg.2017.02.050.
3. Annamma Joy, John F. Sherry Jr, Alladi Venkatesh, Jeff Wang & Ricky Chan (2012) Fast Fashion, Sustainability, and the Ethical Appeal of Luxury Brands, Fashion Theory, 16:3, 273-295, DOI: 10.2752/175174112X13340749707123.
4. Qiu, Jack Linchuan. *Goodbye iSlave: A manifesto for digital abolition*. University of Illinois Press, 2017.
5. Getting everyone to act on climate change, SDG Action, 13 October 2021, https://sdg-action.org/getting-everyone-to-act-on-climate-change/, Accessed: August 15, 2022.
6. Sustainable fashion: Why it matters, and how to identify the winners, J.P. Morgan, April 2021, https://am.jpmorgan.com/kr/en/asset-management/institutional/investment-strategies/sustainable-investing/sustainable-fashion/, Accessed: August 15, 2022.
7. Neumann, H.L., Martinez, L.M. and Martinez, L.F. (2021), "Sustainability efforts in the fast fashion industry: consumer perception, trust and purchase intention", *Sustainability Accounting, Management and Policy Journal*, Vol. 12 No. 3, pp. 571-590. https://doi.org/10.1108/SAMPJ-11-2019-0405.
8. Yongjian Li, Xiukun Zhao, Dan Shi, Xiang Li, "Governance of sustainable supply chains in the fast fashion industry', *European Management Journal*, Volume 32, Issue 5, 2014,
9. McNeill, Lisa, and Rebecca Moore. "Sustainable fashion consumption and the fast fashion conundrum: fashionable consumers and attitudes to sustainability in clothing choice." *International Journal of Consumer Studies* 39, no. 3 (2015): 212-222.
10. M. K. Brewer, "Slow Fashion in a Fast Fashion World: Promoting Sustainability and Responsibility," *Laws*, vol. 8, no. 4, p. 24, Oct. 2019, doi: 10.3390/laws8040024.
11. Statista, Digital Markets, e-Commerce, Fashion, https://www.statista.com/outlook/dmo/ecommerce/fashion/accessories/worldwide, Accessed August 16, 2022.
12. Duygu Turker, and Ceren Altuntas, "Sustainable supply chain management in the fast fashion industry: An analysis of coporate reports", *European Management Journal*, Volume 32, Issue 5, October 2014, Pages 837-849. Accessed: December 28, 2022.
13. Niinimäki, K. (2017). Fashion in a Circular Economy. In: Henninger, C., Alevizou, P., Goworek, H., Ryding, D. (eds) Sustainability in Fashion. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-319-51253-2_8. Accessed: December 28, 2022.
14. Linda Dauriz, Nathalie Remy, and Thomas Tochtermann, A multifaceted future: The jewelry industry in 2020, McKinsey & Company, February 2014, https://www.mckinsey.com/~/media/McKinsey/Industries/Retail/Our%20Insights/A%20multifaceted%20future%20The%20jewelry%20industry%20in%202020/A_multifaceted_future_The_jewelry_industry_in_2020.pdf. Accessed: December 29, 2022
15. Zara's Speedy Apparel Supply Chain Visualized, CB Insights, October 26, 2018, https://www.cbinsights.com/research/zara-apparel-supply-chain/, Accessed: December 29, 2022

16. Martin Geissdoerfer, Paulo Savaget, Nancy M.P.Bocken, Erik Jan Hultink, “The Circular Economy – A new sustainability paradigm?”, *Journal of Cleaner Production*, Volume 143, February 2017, Pages 757-768. Accessed: December 28, 2022.

OPTIMAL LAND CONSERVATION DECISIONS FOR MULTIPLE SPECIES

[redacted]

Abstract. Given an allotment of land divided into parcels, government decision-makers, private developers, and conservation biologists can collaborate to select which parcels to protect, in order to accomplish sustainable ecological goals with various constraints. In this paper, we propose a mixed-integer optimization model that considers the presence of multiple species on these parcels, subject to predator-prey relationships and crowding effects.

1. Introduction

1.1. **Motivation.** Climate change poses a massive threat to the health of humanity [8]. The ramifications of the climate crisis extend beyond increased global temperatures, as necessities like food and water will become more and more scarce. Humans are not the only species that will be affected. At this rate, it's been estimated that one in six species could face extinction [10]. In an interconnected world, it is, therefore, important to incorporate sustainability with a focus on biodiversity into every level of decision making.

Government decision-makers at the municipal, county, state, and federal levels frequently work with private companies, including engineering firms and developers, to determine where development should take place and where "protected areas" must be established. That is, spaces where species are protected from human interference. In this paper, we will be focusing on the design of protected areas.

1.2. **Designing protected areas.** Given an allotment of land divided into parcels, our task is to select which parcels to protect, in order to accomplish ecological goals subject to various financial and economic constraints. These decisions are aided by optimization models. [1, 2]

Protected areas can take on many configurations. Recent work in this area is developing models to ensure particular spatial properties. For instance, some researchers have focused on selecting parcels such that the protected area is connected [3, 6, 4] or contiguous [11, 7]. However, many models that include these additional constraints focus on a single species; incorporating multiple species adds more dimensions to an already computationally expensive problem to solve.

2. Model and Methodologies

In this section, we present the optimization models for making land preservation decisions in the presence of multiple species and budget constraints. We start by introducing our

[redacted]

notation, then present a model where the species do not interact and one where predator-prey relationships are present.

2.1. **Notation.** The notation used in the optimization models and in the rest of the paper are as follows:

Sets:

P	the set of parcels
S	the set of species

Parameters:

w_i	Weight to prioritize species i
$N_i(p)$	Number of individuals of species i observed at parcel p
N_i	Total population of species i across all parcels
$\tilde{N}_i(p)$	Number of individuals of species i that are simulated to be at parcel p in the future
c_p	The cost of preserving parcel p
B	Budget

Decision Variables:

x_p	Binary variable denoting whether or not a parcel is preserved

2.2. **Model without Species Interaction.** Our baseline model assumes that land preservation decisions can be made by taking into account only the number of individuals of each species as currently observed. This means we assume the different species do not interact with each other and that crowding effects do not occur. Another interpretation is that the model takes only the present conditions into consideration, instead of focusing on sustainability.

This baseline model (1) is a knapsack problem, where the objective is to save a weighted combination of species' populations subject to a budget constraint. The weights, w_i, can be chosen to prioritize endangered species or reflect other conservation concerns.

$$
\begin{aligned}
&\text{maximize} && \sum_{i \in S} w_i \sum_{p \in P} N_i(p) x_p \\
&\text{subject to} && \sum_{p \in P} c_p x_p \leq B \\
& && x_p \in \{0,1\} \quad \forall p \in P
\end{aligned} \tag{1}
$$

2.3. **Model with Species Interaction.** The main model we present is the model with species interaction, such as predator-prey relationships. The model itself has the same overall form as (1) with one critical difference. The species' populations $\tilde{N}_i(p)$ are calculated using a simulation that models the numbers of individuals present in each parcel after T time periods. In our numerical testing, we have taken $T = 2000$, which represents steady-state

populations.

$$\begin{aligned} &\text{maximize} && \sum_i w_i \sum_p \tilde{N}_i(p) x_p \\ &\text{subject to} && \sum_p c_p x_p \leq B \\ &&& x_p \in \{0,1\} \forall g \in G \end{aligned} \tag{2}$$

$\tilde{N}_i(p)$ is obtained from $N_i(p)$ using the Gavina et.al.'s model from [5], which adapts the classical Lotka-Volterra equations describing predator-prey relationships to multiple species and takes into account crowding effects. The simulation is described in Algorithm 1.

Algorithm 1: Lotka-Volterra competition with crowding effects

input : $N_i(p)$: Number of individuals of species i observed at parcel p for each i and p

output: $\tilde{N}_i(p)$: Number of individuals of species i that are simulated to be at parcel p in the future for each i and p

foreach *parcel* p **do**
  Initialize the species counts in parcel p, i.e. let $\tilde{N}_i(p) = N_i(p)$ for each i;
  for $t \leftarrow 1$ **to** T **do**
    Let ΔN = (Birth Rate of species i)($\tilde{N}_i(p)$) - (Competition Effect of all species on i) - (Crowding Effect for species i);
    Let $\tilde{N}_i(p) = \tilde{N}_i(p) + \Delta N$.

3. Data

The data was generated using the code provided in Gupta et.al.'s paper [6]. The process is outlined below.

A landscape is an $n \times n$ grid of parcels, where each parcel is a piece of land that can be protected. Each parcel has a value between $[0, 1]$ where a higher value represents a worse habitat for that species. A landscape where better habitat parcels are clustered with each other (and the worse habitat parcels are clustered with each other) has a low habitat fragmentation. A high habitat fragmentation is an absence of these clusters, as the better habitat parcels are dispersed among worse habitat parcels, and vice versa. Based on these landscapes, density of each species are simulated. Given a species population N_i, these individuals are distributed among the landscape using an inhomogeneous point process. This distribution is $N_i(p)$.

Following the model given in [5], we obtained $\tilde{N}_i(p)$ by inputting $N_i(p)$ into Algorithm 1. Due to the scaling in this code, the values outputted are non-integers. To address this, we rounded the output, $\tilde{N}_i(p)$, to the nearest integer.

Habitat preferences were addressed by assigning landscapes and populations to species, which yields unique distributions of each species on a grid. We generated 10,000 10×10 landscapes with random habitat fragmentation levels and the two most and two least fragmented landscapes were selected. For each of these four landscapes, we explored two

population sizes, $N_i = 100$ and $N_i = 250$. This yields 8 different distributions that represent 8 species. Additional details can be found in Table 1.

4. Numerical Testing

With the 8 species generated, we grouped them into sets of size 2 and 5 in order to implement a 2-species reserve and a 5-species reserve. For the 2-species reserves, the sets are $\{S_0, S_1\}, \{S_2, S_3\}, \{S_4, S_5\}$, and $\{S_6, S_7\}$. The 5-species reserve, the sets are $\{S_0, S_1, S_2, S_3, S_4\}$ and $\{S_5, S_6, S_7, S_0, S_1\}$. These give 6 scenarios total: 4 for 2-species reserves and 2 for 5-species reserves.

For each scenario, we varied the budget B from 0 to 100 with a stepsize of 5 and solved (1) with $N_i(p)$ and (2) with $\tilde{N}_i(p)$. We tracked the similarity of each model's solution by counting the number of parcels that had the same protection status, and taking the minimum, mean, and median over all budgets except where budget is 0 and 100. We omitted those values because the solution will always be to protect nothing or protect everything, which is uninteresting for comparison.

The numerical results for all cases can be found in Table 2. In addition, one solution for a 2-species reserve and 5-species reserve are included in Figure 1 and Figure 2. The reserves found using the two models show a high degree of similarity when the objective function weights for the two species are the same. However, when these weights differed, the solutions showed more variation among lower budgets, which is illustrated by Figure 4. A comparison between Figures 1 and 3 show a specific example of the configuration change.

5. Discussion

A drawback to Algorithm 1, is that $\tilde{N}$ needs to be rounded. It is worth exploring alternatives or variations in order to obtain integer solutions. Not only that, but it would be interesting to modify the parameters in Algorithm 1 to explore dynamics that would yield a larger difference between N and $\tilde{N}$.

For future directions, we hope to explore extensions to make the model and numerical testing more realistic. This includes increasing the grid size, expanding methods to obtain N, and investigating different parameters to use for the species weights and parcel costs. A larger grid size would be valuable to pursue because real-world landscapes are typically larger than 10×10. Also, this would allow more possible reserve solutions, thus making the problem more interesting. With regards to the methods to obtain N, estimating the location and movement of a species using the spatial capture-recapture (SCR) model [9] and using these as inputs into our model, as done in Gupta et.al's paper, would provide a more accurate depiction of animal behavior.

References

[1] Diogo Alagador and Jorge Orestes Cerdeira. Operations research applicability in spatial conservation planning. *Journal of Environmental Management*, 315:115172, 2022.

[2] Alain Billionnet. Mathematical optimization ideas for biodiversity conservation. *European Journal of Operational Research*, 231(3):514–534, 2013.

[3] Jennifer K Costanza, James Watling, Ron Sutherland, Curtis Belyea, Bistra Dilkina, Heather Cayton, David Bucklin, Stephanie S Romañach, and Nick M Haddad. Preserving connectivity under climate and land-use change: No one-size-fits-all approach for focal species in similar habitats. *Biological Conservation*, 248:108678, 2020.

[4] Bistra Dilkina, Rachel Houtman, Carla P Gomes, Claire A Montgomery, Kevin S McKelvey, Katherine Kendall, Tabitha A Graves, Richard Bernstein, and Michael K Schwartz. Trade-offs and efficiencies in optimal budget-constrained multispecies corridor networks. *Conservation Biology*, 31(1):192–202, 2017.
[5] Maica Krizna A Gavina, Takeru Tahara, Kei-ichi Tainaka, Hiromu Ito, Satoru Morita, Genki Ichinose, Takuya Okabe, Tatsuya Togashi, Takashi Nagatani, and Jin Yoshimura. Multi-species coexistence in lotka-volterra competitive systems with crowding effects. *Scientific reports*, 8(1):1–8, 2018.
[6] Amrita Gupta, Bistra Dilkina, Dana J Morin, Angela K Fuller, J Andrew Royle, Christopher Sutherland, and Carla P Gomes. Reserve design to optimize functional connectivity and animal density. *Conservation Biology*, 33(5):1023–1034, 2019.
[7] Hayri Önal, Yicheng Wang, Sahan TM Dissanayake, and James D Westervelt. Optimal design of compact and functionally contiguous conservation management areas. *European Journal of Operational Research*, 251(3):957–968, 2016.
[8] Annette Prüss-Üstün, Jennyfer Wolf, Carlos Corvalán, Robert Bos, and Maria Neira. *Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks*. World Health Organization, 2016.
[9] Chris Sutherland, Angela K Fuller, and J Andrew Royle. Modelling non-euclidean movement and landscape connectivity in highly structured ecological networks. *Methods in Ecology and Evolution*, 6(2):169–177, 2015.
[10] Mark C Urban. Accelerating extinction risk from climate change. *Science*, 348(6234):571–573, 2015.
[11] Yicheng Wang, Qiaoling Fang, Sahan TM Dissanayake, and Hayri Önal. Optimizing conservation planning for multiple cohabiting species. *PloS one*, 15(6):e0234968, 2020.

6. Appendix

Species	Fragmentation level	N_i
S_0	highest	100
S_1	2nd highest	100
S_2	highest	250
S_3	2nd highest	250
S_4	lowest	100
S_5	2nd lowest	100
S_6	lowest	250
S_7	2nd lowest	250

Table 1. Details for each species' distribution. Fragmentation level describes the density of habitat quality on a landscape, and N_i is the total population of a species on that landscape.

Case	Min	Average	Median
1	98	99.79	100
2	92	97.26	98
3	92	95.89	96
4	92	95.89	96
5	94	96.95	96
6	94	97.16	98

TABLE 2. These values represent the similarity of solutions given from (1) and (2) over budgets varying from 5 to 95 with stepsizes of 5. For the 100 parcels, the number of parcels with the same protection status is recorded. The minimum, average, and median of this is computed and displayed. Case 1: $\{S_0, S_1\}$, Case 2: $\{S_2, S_3\}$, Case 3: $\{S_4, S_5\}$, Case 4: $\{S_6, S_7\}$, Case 5: $\{S_0, S_1, S_2, S_3, S_4\}$, and Case 6: $\{S_5, S_6, S_7, S_0, S_1\}$ were all conducted.

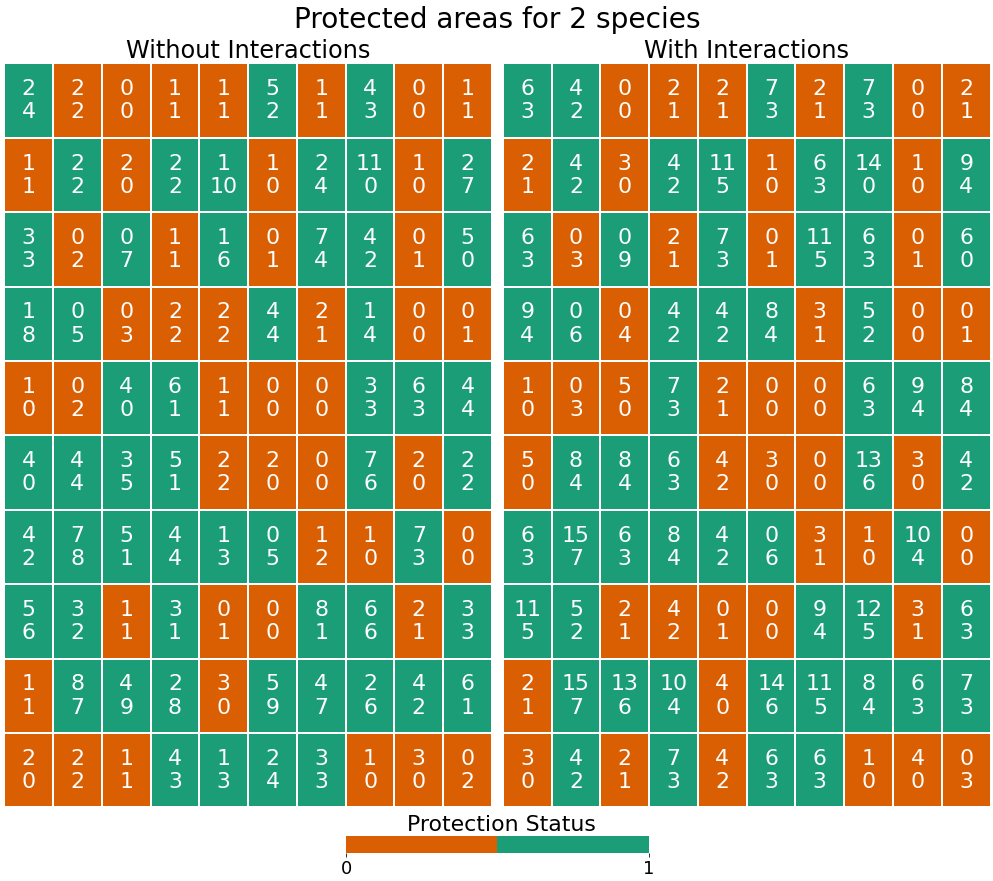


FIGURE 1. Displaying the solution to (1) and (2) for a 2-species reserve using $\{S_2, S_3\}$ when the budget is 55. Each parcel is annotated with its corresponding $N_i(p)$ or $\tilde{N}_i(p)$ for $i = 2, 3$. For each parcel, the value for S_2 is on top and S_3 is on bottom. The color of each parcel represents the preservation decision: green if preserved, orange if not. In this example, 92/100 parcels have the same protection status.

FIGURE 2. Displaying the solution to (1) and (2) for a 5-species reserve using $\{S_0, S_1, S_2, S_3, S_4\}$ when the budget is 55. Each parcel is annotated with its corresponding $N_i(p)$ or $\tilde{N}_i(p)$ for $i = 0, \ldots, 4$. The order of the values in each parcel are top left, top right, middle, bottom left, bottom right, for species S_0, S_1, S_2, S_3, S_4, respectively. The color of each parcel represents the preservation decision: green if preserved, orange if not. In this example, 96/100 parcels have the same protection status.

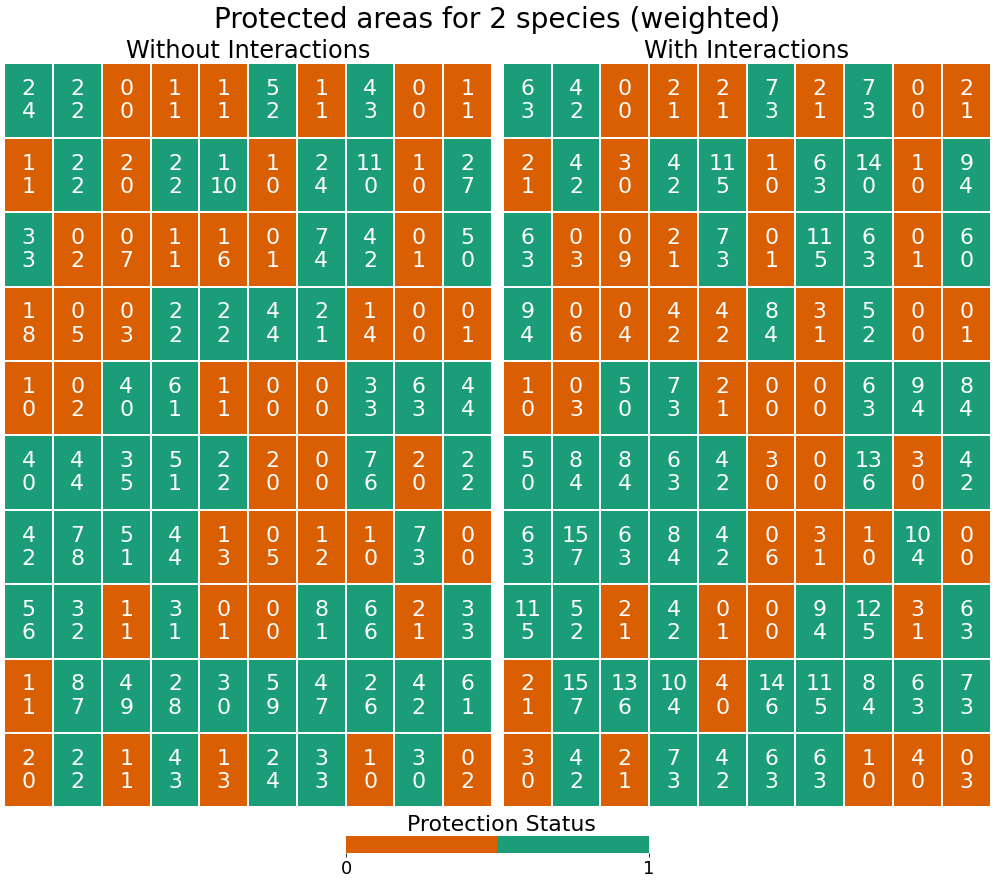


FIGURE 3. This graph depicts the same scenario as Figure 1, but rather than setting $w_2 = w_3$, we used $w_2 = 0.9$ and $w_3 = 0.1$. In this weighted example, 90/100 parcels have the same protection status.

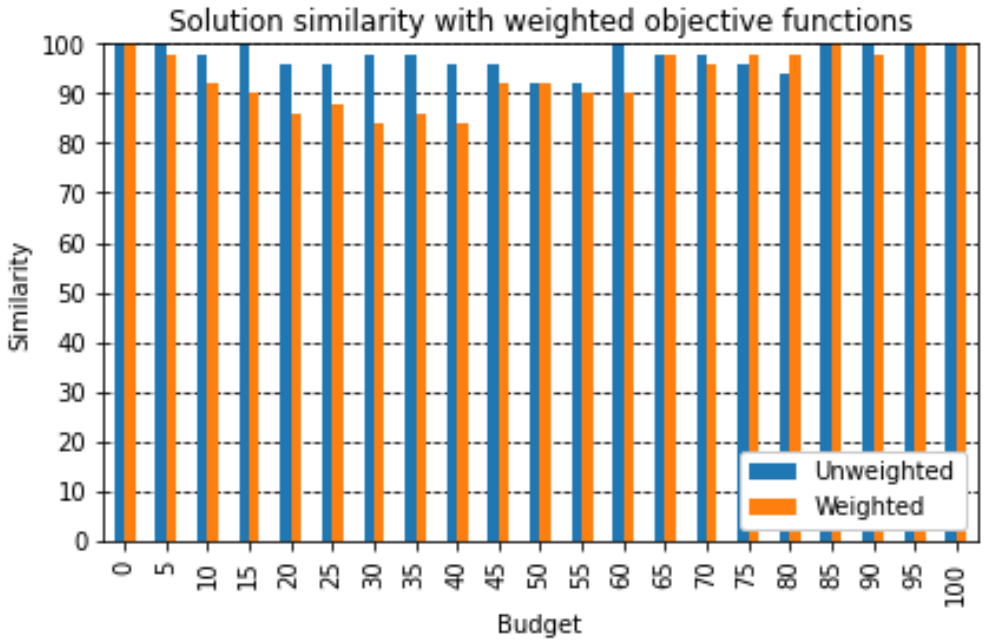


FIGURE 4. The similarity of solutions from (1) and (2) vary based on the objective function weights, w_i. In this example, we used S_2 and S_3 for a 2-species reserve, and solved the models where $w_2 = w_3$ (unweighted) and where $w_2 = 0.9$ and $w_3 = 0.1$ (weighted). For lower budgets, the difference is more apparent.

THE ROLE OF SMALL BUSINESSES IN SUPPLY CHAIN SUSTAINABILITY

ABSTRACT

In today's global economy, supply chain sustainability has become a critical issue for all business enterprises including small businesses. The factors that contribute to sustainability have been well studied. However, there is still a dearth in the current literature that addresses the integration of various factors that will help build sustainable supply chains, especially with respect to small businesses. This paper seeks to fill this gap by developing a holistic approach to sustainable supply chains for small business enterprises.

Keywords: Small Business, Supply Chain, Sustainability, Green Business

INTRODUCTION

Small businesses account for about 44% of U.S. economic activity, and the past few years have been especially challenging as these businesses have navigated the pandemic, financial strain, and supply chain disruptions. A Software Advice survey shows that small and midsize retail businesses (SMBs) are disproportionately impacted by the current supply chain crisis, compared to enterprises that have more resources to weather disruptions. Supply chains must respond to environmental pressures from several sources. Resource availability and regulatory pressures place physical, legal, and economic constraints on supply chain management. Consumer demands and ethical responsibilities of corporations define the desirable behavior within those constraints (Paquette, 2006). As a result of this, supply chain sustainability has become a critical issue for all companies.

Recent innovations in information technology (IT) have positively impacted all aspects of business including the field of supply chain management (SCM). With the emergence of web services, the convergence of telecom and computing is finally reaching maturity in a unified platform for doing business in the 21st century. The widespread availability of highly flexible, functional, and inexpensive information and communication technologies provides us with opportunities for a radical redesign of supply chains (Gunasekaran, 2004). Redesign of supply chain should include a rethinking with sustainability in the background.

This paper is organized as follows. First, we briefly describe the current scenario pertaining to small businesses. This is followed by a brief overview of supply chain sustainability. Business process reengineering relevant to supply chains is discussed next. Concluding remarks form the last section.

SMALL BUSINESSES

Even though the economic importance of small and medium scale enterprises has been known, they were considered comparatively unimportant during the great Internet boom during the 1990s and early 2000s (Passerini, 2012). Use of broadband information technology required extensive investment in technological assets and a long-term access to capital. Such capital requirements were not available to small businesses. Today small businesses can compete and

excel due to continual improvements in Internet technology as well as breakthroughs in cloud computing and mobile connectivity.

Access to capital and an established brand name are the main advantages of large organizations. One of the greatest advantages small businesses have is flexibility. Many small businesses have a single owner who is free to change policies, and technologies (Sadowski, 2002). For example, the owner of a small grocery store can decide to use broadband to create an automatic reorder system with suppliers. Small businesses can offer new services and change internal processes without having to clear a multitude of committees that would exist in a large organization. Cloud computing, and open-source software have brought down the investment requirements and costs. This has resulted in the availability of broadband technologies to small businesses to streamline business processes, grow the customer base, and enlarge existing offerings.

Growth →	*Low*	*High*
Innovation ↓		
High	Constrained	Glamorous
Low	Core	Ambitious

Figure 1. Kirchoff's Typology of Small Firms
[Adapted from Passerini (2012)]

According to Kirchoff's typology (Kim, 2004), small firms can be classified based on two dimensions: innovation and growth as shown in Figure 1. Core firms represent where innovation and growth are low, whereas Glamorous firms have innovation and growth at a high level. Constrained companies have low growth potential but high innovation potential. Ambitious firms have high growth potential but low innovation potential. This paper mostly focuses on Core firms. Mills (2015) classifies small businesses based on types of firms – whether they are sole proprietorships, B2B, etc. – as indicated in Figure 2. Most of the Core small businesses are either sole proprietorships, or local businesses serving consumers and other local businesses. Lower costs of Information Technology (IT) deployment, mobility advantages supported by broadband, and an IT services support system (now directly available as-a-service) can help the more IT conservative small firms (such as 'core' as indicated in Figure 1) to transition to the new mobile apps (Passerini, 2012).

Types of Firms	Number of Firms	Description
Non-Employee Business	23 million	Sole proprietorships
Main Street	4 million	Local businesses
Suppliers	1 million	Suppliers to other businesses
High-Growth	200,000	Fast-growing, innovation-driven businesses

Figure 2. Types of Small Businesses [Adapted from (Mills, 2015)]

There is a lack of knowledge about the potential benefits of information technology and strategies to support small businesses in achieving their business objectives. Small businesses face the challenge that generally they are owner managed and the owner makes all or most of the decisions about the business (Fillis, 2004; Spencer, 2006). Unfortunately, owner-manager's limitations become limitations of the business. Information technology needs to be considered a key player for the small business in reaching its goals. As information technology is perceived to

be expensive by small businesses, they often do not budget for it. The other problem regarding the cost of IT is that small businesses may invest in unnecessarily big solutions due to sales pitches, hype of specific products or market patterns without considering their real need (Grandon, 2004).

With reference to small businesses, applications suitable for blockchain technology can be studied under the following three categories: 'Creating Unbreakable Contracts,' "Safer Data Storage for an affordable Price,' and 'Reduced Complexity in Supply Chains' (TechHQ, 2019). Businesses deal with contracts daily. Blockchain can fill up this part of business by creating smart contracts. As the name suggests, smart contracts are self-executed, coded agreements that deliver guaranteed outcomes if the predetermined conditions are met. The significant difference from paper contract is that smart contract is digitized and that it cannot be tampered with in any way because it is in a blockchain.

SUPPLY CHAIN SUSTAINABILITY

The concept of sustainability is not a new idea. In 1987, the Brundtland Commission (World Commission on Environment and Development) stated that sustainability is development that meets the needs of the present without compromising the ability of future generations to meet their needs. Business sustainability can be defined as the ability to conduct business with a long-term goal of maintaining the prosperity of the economy, environment, and society. This is well described in the triple bottom line (TBL) perspective of sustainability which considers organizational sustainability to include the following three components:

- Environmental Performance (Planet),
- Economic Performance (Profitability), and
- Social Performance (People).

By considering all stakeholders in addition to profits, more sustainable outcomes can be chosen while considering many alternatives. Nearly 70% of the top 250 global companies on the Fortune 500 have adopted TBL reporting (Dao, 2011). The Dow Jones Sustainability Indices (DJSI) launched in 1999, are a family of indices evaluating the sustainability performance of the largest 2,500 companies listed on the Dow Jones Global Total Stock Market Index. Companies vie for a better Dow Jones Sustainability Index.

In a typical manufacturing enterprise, the value chain model consists of the following sequence of activities: inbound logistics, operations, outbound logistics, marketing & sales, and service. The objective is to offer the customer a level of value that exceeds the cost of activities (Chopra, 2007). In addition to the above primary value chain activities, we also need to consider the following supporting activities: firm infrastructure, human resource management, technology development, and procurement. Supply chain management (SCM) can be defined as the combination of art and science of improving the way an enterprise finds the raw components it needs to make a product and delivers it to customers. Viewed from this perspective, SCM is an important part of the overall value chain model. This approach is useful in analyzing the impact of Information and Communications Technologies (ICT) on the manufacturing environment. Sustainability affects all links of the supply chain.

Typically, a supply chain includes all entities and processes involved in fulfilling a customer order. More than one decision maker is involved in managing resources, information, and processes that span several organizations. In a nutshell, supply chain consists of the following stages:

- Sourcing,
- Transformation,
- Delivery,
- Product Use and Recycle.

Sourcing refers to procurement of raw materials, parts, and subassemblies required to produce the final product. A critical part of sustainable supply chain is to have procurement that support environmentally friendly practices. This will be effective when the major entity in the supply chain forces its upstream suppliers to adopt and adapt technology and practices that result in sustainable material sources.

Transformation refers to the conversion of inputs to outputs. Using electricity produced from green alternatives instead of fossil fuels, using recycled paper and plastics where possible and using non-toxic chemicals in their processes are some of the ways to ensure sustainability. The delivery stage includes many operational processes. The choice of facility location such as offshore vs. onshore, close to the customer vs. close to the raw material source can have a significant impact on the greenhouse gas (GHG) emissions. The choice of mode of transportation is another critical factor. Rail and waterways have lower emissions for ton of goods shipped but they are slower than trucks and airplanes which have higher costs and emissions. Inventory management is another area that can impact sustainability. Single period and multi-period inventory models take into consideration stock out and backorder costs but not sustainability criteria.

Consumers have a critical role to play in making supply chain sustainable. Considering products such as cars and computers, a major proportion of emissions comes from consumer's use of products. A key objective of any business should be to make goods that are more energy efficient. A key component of Closed-loop supply chains is the concept of reuse, recycle and return. The objective is to ensure that the product is eventually disassembled, and components reused, re-manufactured or recycled into a source of raw materials.

Elkington (2004) argues that due to the advance of information technology (IT) companies can no longer keep their practices secret from stakeholders. They must report on their sustainability practices to inform them and to serve as a benchmark against competitors. Hassini (2012) lists the following factors as drivers for the adoption of sustainable supply chain practice:

- Market Forces,
- Policy and Regulations,
- Science and Technology,
- Product Development,
- Process Capability,
- Sourcing and Operations,
- Transport and Logistics,
- Marketing and PR, and
- Social Issues.

Market forces factors include consumers, retailers, original equipment manufacturers who may demand products considered environmentally friendly from their suppliers. Financial stakeholders such as mutual funds and pension funds require that the company follow sustainable

practices. It is likely that in the future, access to capital markets may be restricted only to businesses that are deemed to be environmentally friendly. Competition in the marketplace may require a company a company to offer product considered as socially responsible, green, or sustainable. Governments either through legislation or via a regulator requiring that companies adhere to certain environmental standards is an example of how policy and regulations can influence companies to consider sustainability. In case of environmental disasters, governments can retroactively introduce legislation or regulation to curtail some business practices. Klassen and Vachon (2003) found that adoption of ISO 4001 is significantly related to the efforts of companies to invest more in environmental management practices.

The science and technology factor stems from the need to use research and development division to find materials and processes that are not toxic, use less energy or suitable substitutes without compromising use. The product development factor involves using more recycled content, using biodegradable materials or alternative sources of fuels and materials. It also calls for using reverse logistics and design for disassembly. The process capability factor calls for greening the process. This involves using energy efficient machines, fuel efficient transportation, etc. The process of producing the product will have to be environmentally sustainable. In case of returns after the useful life of the product, the supply chain would also have to ensure that the process is capable of absorbing returns and use them back in the production process,

The sourcing and operations factors focus on sourcing of recycled or renewable raw materials, parts or subcomponents and use processes that are more environmentally friendly. Companies like Subaru and Toyota operate zero-waste facilities as a means of reducing costs and helping the environment. They secure zero waste by making sure that no byproduct of their operations ends up in a landfill. The transport and logistics factors direct companies require the use of reverse logistics and closed loop supply chains and reuse, recycle, and return programs.

The marketing and public relations factors pertain to the efforts of companies to create a value proposition to customers. Companies have to create awareness of the practices that makes the product more environmentally friendly and sustainable. *Carbonfund* and *Bullfrogpower* are organizations that provide carbon offsetting and greenhouse gas reduction options to businesses, and organizations. Companies can use certifications from such entities to inform customers that their products are environmentally friendly. The social issues factor focuses more on the existing behavior and practices of companies in relation to the treatment of their labor force, sourcing practices and environmental impact on their communities (Wang and Lin, 2007). Sustainable operations are more concerned with translating laudable aspirations into economically sustainable business practices.

BUSINESS PROCESS REENGINEERING FOR SUSTAINABILITY

Redesign, retooling, and re-orchestrating form the key components of Business Process Reengineering (BPR) that are essential for an organization to focus on the outcome that it needs to achieve. The entire technological, human, and organizational dimensions may be changed in BPR. Information technology plays a major role in business process reengineering as it provides office automation, it allows the business to be conducted in different locations, provides

flexibility in manufacturing, permits quicker delivery to customers, and supports rapid and paperless transactions (Bogdanoiu, 2014).

To make the supply chain sustainable, small businesses should aim to minimize waste. If the enterprise is focusing on selling products, then the supply chain is likely responsible for the bulk of the environmental impact. Small businesses can turn their traditional supply chains into green supply chains by integrating sustainable environmental processes. Some of the ways this can be achieved is listed below:

- Choosing a diverse set of suppliers that match your sustainability standard,
- Choosing and purchasing the correct materials,
- Sustainable product design,
- Sustainable product manufacturing and packaging,
- Sustainable assembling, distribution, and recycling.

The BPR technique implements organizational change based on rapid change, employee empowerment, and training and support by information technology. To implement BPR to an enterprise, the following key actions need to take place:

- Selection of the strategic processes for redesign,
- Simplify new processes – minimize steps – optimize efficiency – modeling,
- Organize a team of employees for each process,
- Organize the workflow – document transfer and control,
- Assign responsibilities and roles for each process,
- Automate processes using information technology,
- Train the process team to efficiently operate the new process,
- Introduce the redesigned process into the new organizational structure.

Although technology can create new or modified business practices at a rapid rate, successful adoption of new best practices must stand up to market forces. Technology and the marketplace are continually reshaping business activities and consequently, business strategies. An organization must continually work towards an alignment that fits into the organization's sustainability strategy and IT strategy. This alignment should improve the likelihood that new initiatives are explicitly linked to areas that are critical to successful business performance, provide a source of competitive advantage. The role of IT should be that of a strategic enabler for competitive success, rather than just an operational supporter. Information technology tools are well suited for develop strategies to create sustainable supply chain functions.

Dao (2011) proposes an integrated sustainability framework. Short term internal strategy emphasizes preventing pollution by optimizing internal operations to reduce cost and have a positive impact on the environment. Strategy also involves creating an organizational culture aimed towards sustainability and improving employee management practices within firms. Payoffs envisaged for activities in this quadrant are reduced costs, increased profitability, and reduced risk. Short term external strategy focuses on improving the extended supply chain to reduce pollution through material and process choices and closed loop supply chain. Strategy also involves extending organizational culture aimed towards addressing sustainability issues affecting both internal and external stakeholders. Payoffs expected for this strategy include reputation, legitimacy, reduced environmental impacts, and increased competitive advantage.

Long term internal strategy involves developing capabilities that enable radical clean technologies and processes that help solve social and environmental issues. Expected payoff for this strategy includes innovation and strategic positioning. Long term external strategy consists of including core sustainability capabilities in all products, processes, and supply chains. This strategy will give an impetus to open new, previously ignored dialogues with stakeholders to solve social issues and locate growth opportunities, thus creating a sustainability vision. Expected payoff would be a growth trajectory. As Paquette (2006) points out, a company that is reactive, flexible, and efficient in execution may operate extremely well at environmental pressure, while a company that is proactive, innovative, and differentiated from competition may best place themselves beyond pressure.

CONCLUSION

Enterprises of all sizes that are parts of supply chain can benefit from using techniques made available by information technology in their efforts to create a sustainable supply chain. Business process engineering has been used successfully in the corporate world. However, optimal use of these techniques focused on sustainability requires domain specific analysis (Watson, 2010). The basic building blocks of IT implementation consist of digitized versions of interactions among various business processes. In this paper, we have presented an approach based on BPR that can identify and categorize the different types of business processes/transactions geared towards sustainability. Restructuring these processes and then automating them in a systematic affords a practical approach to leverage information technology. Monitoring the critical success factors will help in evaluating the success of these measures.

Future work in this area focuses on developing a comprehensive framework that will enable small business entrepreneurs and researchers to point out the potential priority areas that need to be automated first and yield a realistic estimate of resources needed to achieve such transformation. In addition, such an approach will also help in giving a better insight into process restructuring with sustainability as the focus.

REFERENCES

Bogdanoiu, C. (2014). *Business Process Reengineering Method*. http://www.cesmaa.eu/ awards/ Best Student Paper_BogdanoiuCristiana.pdf

Dao, Viet, S., et al, 2011. From Green to Sustainability: Information Technology and an Integrated Sustainability Framework. *Journal of Strategic Information Systems*, 20 (2011) 63-69.

Elkington, J., 2004. Enter the triple bottom line. In: Henriques, A., Richardson, J. (Eds.), The Triple Bottom Line: Does It All Add up? Earthscan, London, pp. 1– 16.

Fillis, I. (2004). Factors Impacting on E-Business Adoption and Development in the Smaller Firm. *International Journal of Entrepreneurial Behaviour and Research,* 10(3), 178-191.

Grandon, E. E. (2004). Electronic Commerce Adoption: An Empirical Study of Small and Medium US Businesses. *Information & Management*, 42(4), 197-216.

Gunasekaran, A., E.W.T. Ngai, 2004. Information Systems in Supply Chain Integration and Management. *European Journal of Operational Research*, 159 (2004) 269-295.

Hassini, Elkafi, J., et al, 2012. A Literature Review and a Case Study of Sustainable Supply Chains with a Focus on Metrics. *International Journal of Production Economics*, 140 (2012) 69-82.

Kim, H. M. (2004). Best Practices in E-Business Process Management: Extending a Re-engineering Framework. *Business Process Management Journal*, 10(1), 17-27.

Klassen, R.D., Vachon, S., 2003. Collaboration and evaluation in the supply chain: the impact on plant-level environmental investment. *Production and Operations Management* 12 (3), 336–352.

Mills, K. (2015). The Four Types of Small Businesses. *Harvard Business Review* (April).

Melville, N., 2010. Information systems innovation for environmental sustainability. *MIS Quarterly* 34 (1), 1–21.

Passerini, K., A. Tarabishy, and K. Patten. (2012). *Information Technology for Small Business*. New York, NY: Springer.

Paquette, J., 2006. The Supply Chain Response to Environmental Pressures, 2006, p.47, available at http://dspace.mit.edu/handle/1721.1/34530

Sadowski, B. M. (2002). Strategic Use of the Internet by Small and Medium-Sized Companies: An Exploratory Study. *Information Economics and Policy*, 14(1), 75-93.

Spencer, A. S., et al. (2006). Schumpeter and New Technology Based Firms, International Enterprise Management Journal, 2: 145-156.

Stonebraker, P.W., Goldhar, J., Nassos, G., 2009. Weak links in the supply chain: measuring fragility and sustainability. *Journal of Manufacturing Technology Management* 20, 161–177.

TechHQ (2019). How can blockchain be used by SMEs? Retrieved on May 26, 2020 from https://techhq.com/2019/08/how-can-blockchain-be-used-by-smes/

Wang, L., Lin, L., 2007. A methodological framework for the triple bottom line accounting and management of industry enterprises. *International Journal of Production Research* 45 (5), 1063–1088.

Watson, R., Boudreau, M., Chen, A., 2010. Information systems and environmentally sustainable development: energy informatics and new directions for the IS community. *MIS Quarterly* 34 (1), 23–38.

Strategy and Organizational Behvior

No Treatment Without a Diagnosis: Government Regulatory Exclusion Influences on Employees Work Environments, their Behaviors, Attitude and Engagement and Perceived Performance in a Healthcare Setting.

ABSTRACT

The CLIA law created an environment of categorization of "in-group" with regard to "out-group" in which some professionals were recognized and others were omitted. Leaving professions like histology technicians (HT), histotechnologist (HTL) and pathologists' assistants (PA) out by omission, and specialties like Clinical Laboratory Scientist (CLS)/Medical Laboratory Technologist (MLT), Pathologist and Cytotechnologist (CT) in. A total of 324 surveys were collected. The HT, PA, and HTL were merged into the out-group, and CLS/MLS, MLTs, Pathologist, and CTs were merged into the in-group. Out-group is associated with decreased performance and higher turnover intention as compared to the in-group. Being in the out-group decreases performance by (β= -.573). Procedural justice (PJ) and job satisfaction (JS) partially mediated the relationship between group class and performance. PJ had an indirect and negative effect of (β= -.3575) and JS had a negative effect (β= -.1840) in the relationship between group class and performance. PJ and JS fully mediated the relationship between group class and turnover intention. Out-group identification was significantly predicted by group class (β = .427), and turnover intention was predicted by out-group identification (β = .246). Out-group had lower levels of professional identification and higher levels of turnover intention. This research shows us the effects that disenfranchisement has had on the "discredited" careers in terms of employees' perception of unfair work environment, lower JS and lower professional identification. It highlights the pivotal roles that social groups play in increasing dissatisfaction and predicting employees' behaviors, feelings of injustice, lower performance, JS, and higher withdrawal.

1: INTRODUCTION

This study explored justice and identity factors affecting AP laboratories units' performance and quality, and histological laboratory staffs' turnover rates, their job satisfaction, professional identification, self-development believes, and skill acquisition and improvement. Also, this paper looked into the over-arching implications for healthcare organizations and stakeholders of current regulatory landscape, analyze how the core services of the AP lab may either contribute or detract from allowing the organization to attain its performance and financial strategic goals. Congruent with other writers' observations, this paper will attempt to generate ideas on how to implement a structural approach to address the current challenges facing this clinical diagnostic field, and bluster organizational performance. Histotechnology is the science dealing with the structure of cells and their formation into tissues and organs, and the profession centers on the detection of abnormal pathology and the treatment for the diseases causing abnormalities, such as cancer.

In order to promote uniform quality and standards among all laboratory testing sites in the United States, congress enacted the Clinical Laboratory Improvement Amendments of 1988 (CLIA), which outlined specific performance indicators (Stull, Hearn, Hancock, Handsfield, and Collins 1998). However, in Anatomic Pathology (AP) laboratories, according to CLIA 88, the technical supervisor is the pathologist, and it excluded by omission other professions working in the pathology field from oversight and the many pre-analytic processes they perform because they were deemed relatively simple, minimal risk procedures. Leaving all the professions such as grossing technicians, histology technicians, histotechnologist and pathologists' assistants that work in AP laboratories outside the coverage of the law. Failing to recognize the importance of

an educated work force in histological practices has created disharmony and contention through the laboratory field.

Several studies like Hurst, Nickel, and Hilborne (1998), Lunz, Castleberry, James, and Stahl (1987), and Stull, et al, (1998) have considered the relationship between lab personnel and the testing quality in the lab. These prior studies found supporting evidence that linked the quality of the test performed as being influenced by the testing personnel academic education, clinical training, work experience, and competency assessment examinations. In particular, Lunz et al, (1987) found that laboratories that employee “only certified medical technologists produce significantly more accurate results on proficiency tests than laboratories that employ only non-certified technologists.”

1.1 Purpose and Significance of the Problem

This study fucuses on a natural field experiment and explores the relationship between the negative feelings associate with the perceptions of exclusion and an unjust work environment and organizational outcomes as it relates to withdrawal, turnover intention, and work performance. Professional identification and the individuals’ feelings and attitudes associated with it has been found to impact the quality of services provided by those individuals (Brott, & Myers, 1999). Physiological stressors such as a negative emotional perception an unjust work environment might be the best predictors of measured organizational outcomes, as negative phycological feelings of justice may manifest themselves through lower levels of performance (Feldman, 1996).

Withing this context, dissatisfaction with the current regulatory framework which outlines those professions that the regulations recognize as "accredited" and those who are not recognized

"discredited", may negatively sway the interactions and allocation of resources between the two groups and hinder the "discredited" groups overall job satisfaction. This dissatisfaction may further manifest itself as a contributing factor towards the error rate in AP diagnostic services. These errors can manifest in the form of poor job attitudes, low employee effort, and substandard performance and quality. Leading to negative patient outcomes and additional unnecessary expenses associated with the cost of screening, diagnostic services, and treatment of cancer and many other ailments.

1.2 Research Question

This study examines the possible unintended outcomes that were generated through the implication of this artificial environment of exclusion. According to social identification theory, people defined their self-identity to some extent with the social groups they are part of (Islam, 2014; Tajfel, 78; Tajfel & Turner, 79). The law created an environment of categorization of “in-group” with regard to an “out-group” in which some professionals were recognized and others were not. Such enforced disenfranchisement might create among the disenfranchised a sense that they are excluded from others. The aim in this study is to address the primary research question which to date has not previously been evaluated in the literature. *What effects does disenfranchisement (debarring despondency), have on the “out-group”, their job satisfaction, intrinsic professional drivers, their self-development believes, skill acquisition and improvement, turnover rates and their perceived units' performance and quality, compared to other laboratory professions who are not disenfranchised?* Of the factors affecting non-pathologist AP professionals, this study will explore the experiences of histological laboratory personnel to assess the relationship between the educational preparation, performance, career expectations and their subsequent retention in the laboratory profession. Research into the environmental

effects on the affected testing personnel professional identity can reveal whether, or how, workers self-identify as professionals, their levels of job satisfaction and the impact on them as individuals, the quality of their work, and their perceived status and loyalties.

The hypothesized relationship depicted in Figure 1 says that the effects of in-group and out-group are mediated by justice, satisfaction, and identification. The research model leverages justice theory as a means to contextualize the perception of those in the out-group of disenfranchisement and debarring despondency, feeling of being ignored, professional tainted, a general lack of justice and unfair environment. The social identity theory model, including professional identity is also used to depict the intrinsic drivers to transcend barriers and fill growth needs through service to others, coping by focusing on the greater impact on patients through their work, deriving extrinsic job satisfaction. Pratt et al, (2006) found that behavior in the workplace has been found to depend on how professionals view their work identity, it can also have a profound impact on the overall quality of products and services provided by such individuals.

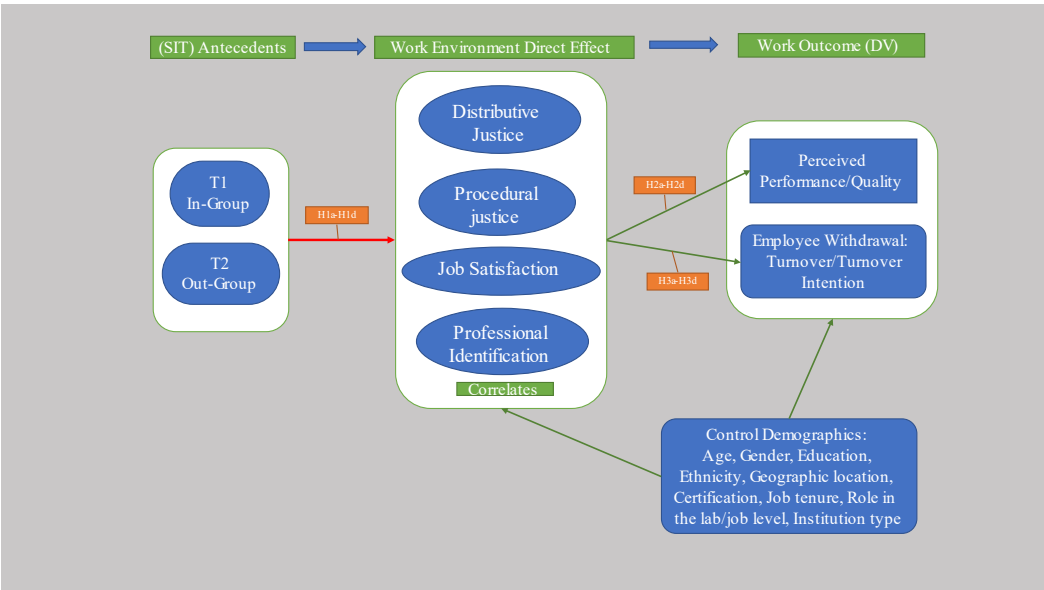


Figure 1. Research Model

1.3 Research Setting

This study brings attention to an issue that has been neglected and aims to highlight the larger policy ramifications and its effects on the rising cost of healthcare in the United States. Thus, this natural occurring field experiment research setting that is being driven by regulatory policy provides a unique opportunity to examine the effects that a perceived unfair work environment has on employee attitudes and work outcomes. Those "out-group" professions were shackled by bureaucratic prescriptions about how to work and are not free to operate as they see fit. They are further barred from climbing up the ladder by regulatory exclusion of their profession to serve as technical supervisors within their fields.

Optimal patient outcomes are a priority for health providers across the nation. Developing high quality levels of care thought increase performance and better quality of products and services by AP lab directly relates to the total resource used and cost of treatment to the hospitals and thus has important strategic implications as time and resources spent due to poor quality pre-analytical processes may cut into the available capacity for performance and thus have a high opportunity cost in terms of forgone revenue and increased liability (Pisano, Bohmer, and Edmondson, 2001). Furthermore, some types of pre-analytical errors in processes are irreversible. As such, no additional compensation is provided to the hospitals for going back and reworking a sample in the AP laboratory due to inadequate quality lab products, causing delays or inaccurate and inconclusive diagnostic services and some of these additional costs are borne by the hospital. This study also is in line with professional organizations and industry recommendations for regulatory agencies to include the histologic professions that perform routine and high complexity pre-analytical processes as a separate category under (CLIA).

1.4 Surgical Pathology Defects and the Role of HP in Error Prevention

The World Health Organization (WHO) estimated that for 2011, cancer topped all coronary heart disease or all strokes as the number one cause of deaths (World Health Organization, 2015). With growth estimates to over 20 million new cancer cases expected annually as early as 2025 (Bray, 2014), and 28 million new cancer cases by 2040 (Sung, Ferlay, Siegel, Laversanne, Soerjomataram, Jemal, & Bray, 2021). Furthermore, a statistical brief by Soni (2015) under The Agency for Healthcare Research and Quality's (AHRQ) that analyzed data from the Household Component of the Medical Expenditure Panel Survey (MEPS) showed that, historically, cancer has been one of the top five most costly medical conditions to treat in the United States. The United States national patient economic burden that has been estimated to be associated with the various stages of cancer care was approximately "12 percent" of the national medical care spending "$21.09 billion" in 2019 (Yabroff, Mariotto, Tangka, Zhao, Islami, Sung… & Ward, 2021). This $21 billion figure was further broken down into patient out-of-pocket costs and patient time costs "$16.22 billion and $4.87 billion" respectively (Yabroff, et al., 2021).

The cost of healthcare services is quite different from the patient's out of pocket cost to the amount that providers incur to deliver that service (Arora, Moriates, & Shah, 2015). The economic burden of pathology diagnostic services on healthcare systems, patients, insurance, and other payers can be substantial. This study aims to provide some perspective on a link in the chain that has not been studied in prior research, which could potentially be of significant impact in contributing to healthcare costs incurred in the diagnostic process of many ailments such as cancer. Although there has been ample research that has led to the development of numerous advancements in testing methodologies, healthcare providers and oncologist should be made aware that there are significant differences between how laboratories operationalize those tests

and the qualifications of staff performing them, which can have a significant impact on patient safety.

Some studies have proposed dividing the types of errors in anatomic pathology into four categories, misinterpretations, misidentifications, specimen defects, and other defects (Meier, Zarbo, Varney, Bonsal, Schultz, Vrbin, ... & Raab, 2008; Meier, Varney, & Zarbo, 2011). Zarbo, Meier, and Raab (2015) further attributed human interactions as the usual source of errors in anatomic pathology. A study in 2014 that looked at 73 institutions, and the rates of surgical pathology defects, found that 14.6% misinterpretations such as false-positive or false-negative, 13.3% of errors misidentifications, 13.7% of specimens had defects, and 58.4% of specimens had other report defects (Volmar, et al., 2014). Their study further found that that most defects, about 47.4 % were often caught internally by a pathologist internally, and that clinicians detected about 22% of defects (Volmar, et al., 2014). Other studies that have looked at the rates of inaccurate diagnoses have also noted wide variance in the percentage of errors reported from 1.3% (Epstein, Walsh, & Sanfilippo, 1996), 5.8% (Abt, Abt, & Olt, 1995), approximately 7%–9% (Peck, Moffat, Latham, & Badrick, 2018), to 16% (Tsung, 2004). The high error rates in anatomic pathology represent not only a risk of inaccurate diagnosis but also fertile grounds for performance and quality improvement which can lead to reduced cost for stakeholders involved.

2: BACKGROUND AND LITERATURE REVIEW

2.1 Social Identity

2.1.1 Professional Identification

Social identity has further branched as a useful framework for understanding professional identity as it has server to provide the context needed for professionals to come together and

interact, which serves as the basis for complex identities to evolve (Hogg, & Williams, 2000). While a person might belong to several distinct groups, research has found that professional group associations often emerge as the one of the most significant in that person's life (Adams, Hean, Sturgis, & Clark, 2006; Turner, 1999). These social groups involve the categorization of "in-group" with regard to "out-group" and the individuals' tendencies to view their group in a positive light as opposed to those in the outgroup with a negative bias. As such, the current regulatory framework which outlines those professions that the regulations recognized as "accredited" or "in-group" and those who were not recognized by virtue of omission as "discredited" or "out-group" might have help produce or influence negative organizational outcomes, which may be manifested through negative feeling of justice, lower performance and quality of services and products by those in the "out-group" (Feldman, 1996).

An unclear professional identity can have many tangible and tacit effects on the individuals that are part of that particular group. It can affect the persons pride and their stability (Nelson & Jackson, 2003), and can even impact the different work boundaries, their developmental opportunities, and their ability to fulfill their roles (Schoen, 1989). When professional identity is not clear, it can also negatively impact the overall satisfaction for workers within a profession, affecting the services provided and increasing dissatisfaction or resentment (Remley & Herlihy, 2005).

2.2 Job Satisfaction

Several papers have examined how job satisfaction might be an antecedent of work-related behaviors such as job performance (Iaffaldano & Muchinsky, 1985; Judge, Thoresen, Bono, & Patton, 2001; Petty, McGee, & Cavender, 1984), and turnover (Tett, & Meyer, 1993). Job satisfaction has generally been described as being composed of an affective or emotional

component about his or her job (Brief, 1998; Brief & Roberson,1989; Organ & Near, 1985). Cammann, et al, (1979) hypothesized and empirically showed that works stressors, which were described as by Jex, Beehr, and Roberts, (1992) as any aspect of a person's work environment that may have the potential to cause mental of physical illness, are an antecedent that is related to that individuals' levels of job satisfaction. Other types of work stressors include role ambiguity and role conflict (Fisher & Gitelson, 1983; Jackson & Schuler, 1985), and organizational constraints (Spector & Jex, 1998). Work stressors were further characterized as having a negative association with job satisfaction (Cammann, et al, 1979). Prior research and theory found that in-role performance is related as a consequence of an individual's level of satisfaction (Iaffaldano, & Muchinsky, 1985; Judge, & Bono, 2001; Petty, McGee, & Cavender, 1984). Furthermore, the relationship link between these variables has been hypothesized to operate via effect on the individual's motivation (Strauss, 1968).

2.3 Altruistic Role Repurposing

Is an emergent category that serves as a vehicle to overcome policy influenced sociological and hierarchical stigma. Prior research has found that as job satisfaction increases, the intent to leave decreases (Chalykoff, & Kochan,1989) and that job satisfaction has been found to be related to both organizational citizenship behavior (Bateman, & Organ, 1983; Smith, Organ, & Near, 1983) and workplace justice (Moorman, 1991). As such OCB was defined by Organ (1988) as extra-role behaviors that are discretionary and are not related to the employees' formal work or reward system. As the excluded careers may express their desire for respect and recognition from the other "accredited" professions, they may explore alternative ways to realizing their personal potentials and finding meaning that transcends their sense of self through service to others.

2.2 Organizational Justice

2.2.1 Fairness perception

The organizational justice literature has largely focused on the relationship between fairness perception and outcomes such as outcome satisfaction, job satisfaction, organizational commitment, trust, evaluation of authority, organizational citizenship behavior, withdrawal, negative reactions, and performance (Colquitt, 2001a; Suliman, 2007; Fernandes & Awamleh, 2006; Cropanzano, et al., 2001; Moorman, 1991). The outcome of job withdrawal has often been associated with behaviors such as absenteeism, turnover, and neglect. Withdrawal has been examined as an outcome that may occur thorough individual's reasoned evaluation of the organization as whole in reaction to an unsatisfactory outcome or by the reaction to poor interpersonal treatment by a leadership figure (Colquitt, 2001a). Borman (1991) defined performance outcomes as a reflection of the members' contributions to the organization's goals. This research explores how the members of a career field have reacted based on their fairness perception of fair or unfair treatment.

The Colquitt, Conlon, Wesson, Porter, & Ng, (2001a) paper on organizational justice research noted justice to be a social construct. The question of the individual's fairness perception is derived from past experiences or antecedents and consequences and linking these objective facets of decision making to subjective perceptions of what might be considered fair (Colquitt, 2001a). They further categorize the types of subjective perceptions into two categories, first is "fairness of outcome distributions", or also referred to as distributive justice (Adams, 1965; Deutisch, 1975; Homans, 1961; and Leventhal, 1976), and second is the "fairness of procedures used to determine the outcome distributions" also referred to as procedural justice (Colquitt, 2001a, p 425; Leventhal, 1980; Leventhal, Karuza, & Fry, 1980; Thibaut and Walker, 1975).

The current laws have created a framework that underlines the interactions between those professions that the regulations recognize are "accredited" and those who are not recognized "discredited". Some adverse behaviors that may fallow and individuals' perception of negative assessment of distributive justice or relative equity as compared to other employees in a similar field might result in reduced job performance, and withdrawal behaviors (Folger, & Cropanzano, 1998).

The excluded career fields sociological and psychological views that they are in an environment where their professional voice does not hold the same value as other professions that were "accredited". In this environment, the excluded careers may experience feelings that they are treated as second class profession, experiencing professional prejudice and exclusion. As a consequence of this regulatory environment that promotes a hierarchically based status that is based on policy characteristics between the different professions, the "discredited" professionals have a sense of devalue professional identity, such that those "accredited" professions feel that their knowledge and actions are tainted as substandard. Furthermore, policy practices such as this serve to lower the "discredited" professions level of autonomy and opportunity to contribute to decision and lower their job satisfaction. Regulatory exclusions are hypothesized to elicit psychological reactance within this context (Brehm, 1972), and such reactance can manifest itself in the form of poor job-attitudes and low employee effort.

Both research and theory from Colquitt, (2001a) and Cohen-Charash and Spector (2001) found that there is a correlation between job satisfaction and organizational justice. If an employee perceived the outcomes or the process by which outcome allocation decisions were made to be unfair, they are likely to reciprocate through lower performance or withdrawal behaviors. Participants with negative feelings of justice are likely to report lower levels of job satisfaction,

with distributive and procedural justice being the best possible predictors of job satisfaction. Furthermore, the affective component of job satisfaction (Brief, 1998), which is emotionally oriented in nature, psychological stressors such as frustration, burnout, emotional exhaustion should yield a particularly strong relationship between organizational justice variable and job satisfaction.

Hypothesis 1a. People in the in-group will have a higher perception of procedural justice than those in the out-group.

Hypothesis 1b. People in the in-group will have a higher perception of distributive justice than those in the out-group.

Hypothesis 1c. People in the in-group will have a higher perception of job satisfaction than those in the out-group.

Hypothesis 1d. People in the in-group will have a higher perception of their professional identification than those in the out-group.

2.3 Performance effects

Prior studies have found team identification is positively linked to innovation and performance (Paulsen, Callan, Ayoko, and Saunders, 2013). Schneider, White and Paul (1998) also found that customer and employee assessment ratings of service quality and performance were related. Other researchers have also found that measures of perceived unit performance to positively correlated with objective measures of organizational performance (Dollinger, & Golden, 1992). Brott and Myers (1999) found that "professional identity development is not only crucial to one's individual growth as a professional, but it can also impact the services provided to those with whom professionals come into contact."

2.4 Learning Orientation

Learning orientation is relevant in this regard as it further focuses the individuals on the acquisition of new knowledge and the development of "deep-processing strategies" which in turn further facilitates the mastery of challenging tasks (Elliot & McGregor, 2001). Alongside these theories, Giddens (1984) theory of structuration considers the extent to which the individual as agent can construct an identity as opposed to being curtailed by the sociological environments that they might arbitrarily find themselves in.

The concept of individual professional identity has been linked with encompassing not only personal work values, but also changes in perception in regard to one's professional identity can further impact the workers level of KSA, growth as a person, success and improvement at work, and imagination and innovation (Gazzola & Smith, 2007). This sociological and regulatory climate further exacerbates the disparities in education level and counteracts the original desire for the regulations for an educated work force as it is positively linked to overall improvements in testing quality. Failing to recognize the established accredited degree programs, culminating in an Associates/Baccalaureate/Masters' degree, and certification through Board of Registry of the American Society of Clinical Pathologists (BOP/ASCP) can have a negative impact on the workers growth, their KSA, and limit their performance capabilities. As such, justice, satisfaction, and identification constructs are hypothesized to fully mediate the relationship between group classification and the outcome variables.

> *Hypothesis 2a. Perceptions of disenfranchisement in distributive justice of the out-group will be associated with decreased performance as compared to the in-group.*
>
> *Hypothesis 2b. Perceptions of disenfranchisement in procedural justice of the out-group will be associated with decreased performance as compared to the in-group.*
>
> *Hypothesis 2c. Perceptions of disenfranchisement of the out-group will be associated with decreased job satisfaction and lower employee performance as compared to the in-group.*

Hypothesis 2d. Perceptions of disenfranchisement of the out-group will be associated with decreased professional identification and lower employee performance as compared to the in-group.

2.5 Withdrawal

Job Satisfaction has been found to be both theoretically and empirically linked to several forms of employee withdrawal behavior, including absenteeism (Farrell, & Stamm, 1988) and turnover intention and actual turnover (Tett, & Meyer, 1993). Prior studies have used turnover intention as a central outcome variable in the study of employee attitudes and behaviors as an outcome of job dissatisfaction and poor working conditions (e.g., Cammann, et al, 1979). Dissatisfied histology staff can't easily leave their career field without significant reinvestment in continuing education to transition towards other laboratory sciences which may not be of interest to them.

Decisions to leave their career field are hypothesized to be driven by their perception of unfairness in pay, upward mobility, input ratio to outcomes and lack of process, decision control. Some researchers have also found that the relationship between turnover intention and turnover is moderated by perceived job mobility and opportunity (Carsten, & Spector, 1987; Wheeler, Gallagher, Brouer, & Sablynski, 2007). The relationship between job satisfaction scales in this paper and turnover intention is expected to be strong, as withdrawal behaviors are a type of lower employee motivation and strategy that they can leverage to avoid unpleasant or dissatisfying work (Hanisch, & Hulin, 1991). In this study, the null hypothesis states that there is no difference between strength of the Perceived Performance and Quality, Turnover, Job Satisfaction, Professional Identification, and Learning Orientation of the out-group as compared to those in the in-group.

Hypothesis 3a. Perceptions of disenfranchisement in distributive justice of the out-group will be associated with increased employee withdrawal as compared to the in-group.

Hypothesis 3b. Perceptions of disenfranchisement in procedural justice of the out-group will be associated with increased employee withdrawal as compared to the in-group.

Hypothesis 3c. Perceptions of disenfranchisement of the out-group will be associated with decreased job satisfaction and increased employee withdrawal as compared to the in-group.

Hypothesis 3d. Perceptions of disenfranchisement of the out-group will be associated with decreased professional identification and increased employee withdrawal as compared to the in-group.

3: RESEARCH METHODOLOGY

3.1 Data Collection and Analysis

This study was performed by conducting a field and online survey that was launched on April 14th, 2022, through Jun 26th, 2022. It is a field study by virtue of the population being split into two treatment groups by prior licensing actions of the government without any connection to the data collection. This study utilized prior used and validated scales in peer reviewed articles to cross reference and determine appropriate wording. A total of 324 surveys were collected, however 59 were discarded due to incomplete or missing data. The data was first validated by analyzing the measurement model and continuing with structural model. This step process started by first considering the reliability and psychometric validity of the measurement model, then in the second step, continuing with an assessment of the structural model in where the hypotheses are tested as well (Hair, 1995).

3.2 Measurement Model

A principal components analysis (PCA) was first performed. Factors that cross-loaded or did not load as expected were dropped one at a time to improve the factorial validity, the coefficient alpha was calculated on the reduced set to determine the degree of interrelatedness of items. To overcome the concern of common method bias in the survey design, several reverse scored items

were incorporated in the principal constructs to reduce acquiescence problems. Data was also verified to rule out multicollinearity, heteroscedasticity, and influential outliers amongst the participants beliefs, by checking the statistical analyses for factors such as the variance inflation factors. Table 1 shows the factor loading model. The finalized items, their associated constructs, descriptive statistics, factors loading, and Cronbach's alphas can be seen in Table 2.

Table 1: Factor Loadings for Validated Measurement Model

Construct	Item	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9
Procedural Justice Measure Items	Q7_PJ1	**0.822**	-0.002	0.113	0.077	0.041	-0.046	0.086	-0.043	0.097
	Q7_PJ2	**0.831**	0.047	0.151	0.085	0.075	-0.120	0.095	0.022	0.111
	Q7_PJ3	**0.593**	0.068	0.462	0.070	0.117	-0.091	0.105	0.104	0.045
	Q7_PJ4	**0.731**	-0.018	0.248	0.071	0.067	-0.157	0.169	0.030	0.097
	Q7_PJ5	**0.702**	-0.025	0.284	0.080	0.077	0.019	0.127	0.010	-0.093
	Q7_PJ6	**0.811**	0.066	0.132	0.060	0.093	-0.060	0.101	-0.018	0.122
	Q7_PJ7	**0.746**	0.051	0.181	0.053	0.227	-0.039	0.110	-0.075	-0.030
Distributive Justice Measure Items	Q8_DJ1	0.080	-0.029	0.133	**0.912**	0.123	-0.103	0.103	-0.009	0.034
	Q8_DJ2	0.088	-0.027	0.079	**0.951**	0.064	-0.124	0.097	0.037	0.071
	Q8_DJ3	0.097	-0.035	0.053	**0.937**	0.072	-0.121	0.113	0.080	0.109
	Q8_DJ4	0.145	-0.009	0.089	**0.913**	0.082	-0.081	0.045	0.007	0.095
Perceived Job Satisfaction	Q9_JSS1_R	-0.063	-0.044	0.028	-0.110	-0.053	-0.141	0.078	-0.047	**0.788**
	Q9_JSS3	0.237	0.103	0.218	0.147	0.086	-0.181	0.102	0.129	**0.689**
	Q9_JSS5	0.079	-0.066	0.134	0.140	0.106	-0.015	-0.040	0.098	**0.715**
	Q9_JSS6	0.139	-0.010	0.114	0.230	0.012	-0.299	0.177	0.066	**0.732**
Perceived Professional Identification	Q10_Identity1	0.294	-0.036	-0.081	0.105	**0.696**	-0.172	0.133	-0.084	0.028
	Q10_Identity2	0.255	0.030	0.020	0.106	**0.746**	-0.064	0.067	0.011	0.156
	Q10_Identity3	0.007	-0.062	-0.086	0.062	-0.174	0.048	-0.105	**0.868**	0.098
	Q10_Identity4	0.020	-0.092	-0.020	-0.024	-0.246	0.126	-0.002	**0.864**	0.035
	Q10_Identity5	-0.020	-0.026	0.005	0.079	-0.139	0.021	-0.167	**0.869**	0.065
	Q10_Identity6	0.138	0.173	0.096	0.060	**0.624**	-0.189	0.001	-0.362	0.014
	Q10_Identity7	0.074	0.042	0.193	0.078	**0.808**	-0.116	0.217	-0.214	0.017
	Q10_Identity8	0.009	0.244	0.152	0.043	**0.712**	-0.141	0.241	-0.146	-0.084
	Q10_Identity9	-0.019	0.168	0.267	0.047	**0.728**	-0.017	0.198	-0.136	0.032
Altruism	Q11_Impact1	0.203	0.071	0.121	0.051	0.093	-0.134	**0.767**	-0.096	0.079
	Q11_Impact2	0.135	0.034	0.092	0.127	0.098	-0.099	**0.781**	-0.083	0.015
	Q11_Impact3	0.101	0.191	0.022	0.033	0.092	0.051	**0.750**	-0.042	0.063
	Q11_Impact4	0.045	0.062	-0.031	0.044	0.195	-0.080	**0.738**	0.037	0.015
	Q11_Impact5	0.232	0.075	0.119	0.124	0.204	-0.038	**0.681**	-0.141	0.114
Perceived Performance	Q12_Quality1	0.227	0.004	**0.736**	0.083	0.122	-0.201	0.093	0.067	0.221
	Q12_Quality2	0.205	-0.035	**0.663**	-0.007	0.076	-0.129	0.166	-0.021	0.050
	Q12_Quality3	0.249	0.001	**0.755**	0.125	0.153	-0.142	0.032	0.095	0.065
	Q12_Quality4	0.249	0.041	**0.669**	0.152	0.135	-0.193	0.073	0.040	0.036
	Q12_Quality5_R	0.128	-0.001	**0.724**	-0.035	0.014	0.019	-0.115	-0.253	0.032
	Q12_Quality6_R	0.175	-0.083	**0.714**	0.107	0.038	-0.089	0.073	-0.083	0.130
Turnover Intention	Q13_Quitting1	-0.137	-0.049	-0.213	-0.036	-0.063	**0.794**	-0.117	0.050	-0.123
	Q13_Quitting2	-0.075	0.099	-0.182	-0.086	-0.118	**0.853**	-0.117	0.028	-0.099
	Q13_Quitting3	-0.058	-0.011	-0.212	-0.196	-0.047	**0.802**	-0.163	-0.020	-0.240
	Q13_Quitting4	0.012	-0.085	-0.164	-0.134	-0.179	**0.742**	-0.086	0.198	-0.160
	Q13_Quitting5	-0.162	0.188	0.056	-0.049	-0.172	**0.580**	0.145	0.014	-0.001
Learning Goal Orientation	Q14_Improving1	0.026	**0.563**	0.089	-0.039	0.286	-0.042	0.133	0.055	-0.023
	Q14_Improving2	0.008	**0.880**	0.005	-0.001	0.069	0.048	0.118	-0.108	-0.069
	Q14_Improving3	-0.021	**0.889**	0.008	-0.043	0.057	0.034	0.055	-0.079	-0.096
	Q14_Improving4	0.006	**0.903**	-0.011	0.001	0.030	-0.008	0.061	-0.089	-0.053
	Q14_Improving5	0.053	**0.675**	0.010	0.001	-0.003	-0.016	0.115	0.023	0.087
	Q14_Improving6	0.057	**0.756**	-0.153	-0.017	0.075	0.106	-0.064	-0.013	0.074

Table 2. Constructs and items with statistics, factor loading and Cronbach's Alpha

Construct	Item	Mean	Standard Deviation	Factor Loading	Cronbach's Alpha
Procedural Justice Measure Items	Q7_PJ1	5.008	1.807	0.822	0.905
	Q7_PJ2	4.772	1.857	0.831	
	Q7_PJ3	4.581	1.770	0.593	
	Q7_PJ4	5.000	1.609	0.731	
	Q7_PJ5	5.333	1.505	0.702	
	Q7_PJ6	5.093	1.584	0.811	
	Q7_PJ7	5.589	1.422	0.746	
Distributive Justice Measure Items	Q8_DJ1	4.480	1.880	0.912	0.969
	Q8_DJ2	4.484	1.825	0.951	
	Q8_DJ3	4.336	1.866	0.937	
	Q8_DJ4	4.307	2.029	0.913	
Perceived Job Satisfaction	Q9_JSS1_R	3.153	1.720	0.788	0.793
	Q9_JSS3	3.756	1.641	0.689	
	Q9_JSS5	3.701	1.427	0.715	
	Q9_JSS6	3.813	1.679	0.732	
Out-Group Perceived Professional Identification	Q10_Identity3	2.029	1.258	0.868	0.900
	Q10_Identity4	1.857	1.159	0.864	
	Q10_Identity5	1.827	1.139	0.869	
In-Group Perceived Professional Identification	Q10_Identity1	5.556	1.349	0.696	0.872
	Q10_Identity2	5.038	1.557	0.746	
	Q10_Identity6	6.143	1.021	0.624	
	Q10_Identity7	5.898	1.055	0.808	
	Q10_Identity8	5.975	1.094	0.712	
	Q10_Identity9	5.595	1.291	0.728	
Altruism	Q11_Impact1	5.335	1.397	0.767	0.845
	Q11_Impact2	5.436	1.414	0.781	
	Q11_Impact3	5.485	1.310	0.750	
	Q11_Impact4	5.321	1.408	0.738	
	Q11_Impact5	5.768	1.292	0.681	
Perceived Performance	Q12_Quality1	4.316	1.686	0.736	0.867
	Q12_Quality2	5.397	1.307	0.663	
	Q12_Quality3	4.697	1.664	0.755	
	Q12_Quality4	5.017	1.534	0.669	
	Q12_Quality5_R	4.141	1.818	0.724	
	Q12_Quality6_R	3.842	1.938	0.714	
Turnover Intention	Q13_Quitting1	3.137	2.019	0.794	0.868
	Q13_Quitting2	3.352	2.029	0.853	
	Q13_Quitting3	3.854	2.128	0.802	
	Q13_Quitting4	3.258	1.979	0.742	
	Q13_Quitting5	4.185	1.854	0.580	
Learning Goal Orientation	Q14_Improving1	5.405	1.364	0.563	0.863
	Q14_Improving2	5.948	0.966	0.880	
	Q14_Improving3	5.948	1.018	0.889	
	Q14_Improving4	5.875	1.035	0.903	
	Q14_Improving5	5.346	1.313	0.675	
	Q14_Improving6	5.793	1.040	0.756	

3.3 Sociodemographic Characteristics

Upon examination of our sample population shown in Table 3. Most of the respondents were 35-44 years old (29%), followed by 45-54 years old (27%). Caucasians and people that identified

themselves as Asian accounted for the largest race/ethnicity groups (52% and 22%), respectively. Gender distribution consisted of 73% females, 25% males, and 1% non-binary. 99% declared that they had post high-school education with 9% having some college courses, 17% had 2-year degrees, 31% had 4-year degrees, 34% had professional degrees, and 8% declared that they had a doctoral degree. 54% of participants reported working in non-profit hospitals or healthcare systems, and 57% reported the current role in their organization as team members, 22% were lead/supervisors and 12% were management.

Table 3. Respondent Demographics

Age	N	%
18 - 24	1	0%
25 - 29	19	8%
30 - 34	42	18%
35 - 44	65	29%
45 - 54	61	27%
55 - 64	33	14%
65 - 74	5	2%
Prefer not to disclose	2	1%

Gender	N	%
Male	56	25%
Female	166	73%
Non-binary / third gender	3	1%
Prefer not to say	2	1%

Institution	N	%
Non-profit hospital or healthcare system	142	54%
For-profit hospital or healthcare system	45	17%
Private, independent laboratory	29	11%
Reference laboratory	12	5%
VA hospital	6	2%
Research laboratory	14	5%
Veterinary laboratory	5	2%
Other. Please specify	6	2%
Military	6	2%

Role	N	%
Team member	150	57%
Lead/Supervisor	58	22%
Manager	32	12%
Director	23	9%
Other. Please specify	2	1%

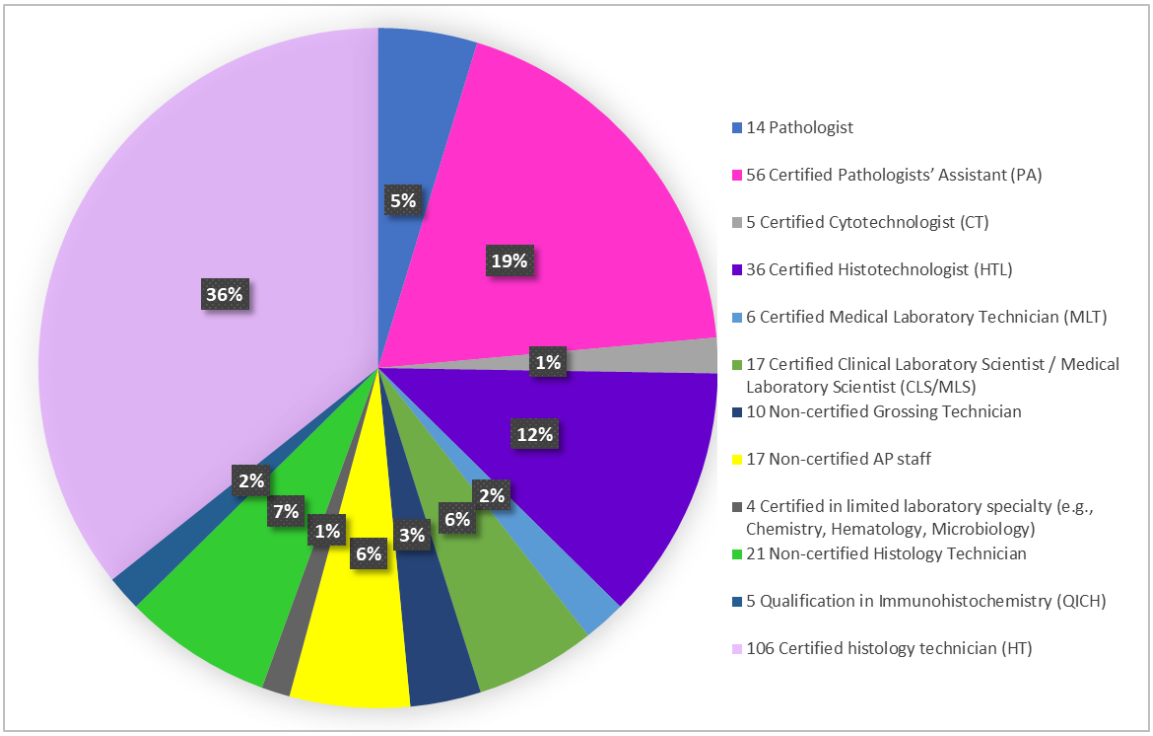
Ethnicity	N	%
White or of European descent	119	52%
Black or African American	12	5%
American Indian or Alaska Native	1	0%
Asian	51	22%
Native Hawaiian or Pacific Islander	3	1%
Mexican or Central American (Latin X) descent	20	9%
Other	11	5%
Prefer not to disclose.	11	5%

Certification	N	%
Non-certified AP staff	17	6%
Non-certified Histology Technician	21	7%
Certified Histology Technician (HT)	106	36%
Certified Histotechnologist (HTL)	36	12%
Certified Pathologists' Assistant (PA)	56	18%
Non-certified Grossing Technician	10	4%
Qualification in Immunohistochemistry (QICH)	5	1%
Certified Cytotechnologist (CT)	5	2%
Certified Medical Laboratory Technician (MLT)	6	2%
Certified Clinical Laboratory Scientist / Medical Laboratory Scientist (CLS/MLS)	17	6%
Certified in limited laboratory specialty (e.g., Chemistry, Hematology, Microbiology)	4	1%
Pathologist	14	5%

Education	N	%
High school graduate	3	1%
Some college	20	9%
2-year degree	39	17%
4-year degree	70	31%
Professional degree	78	34%
Doctorate	18	8%

The following other pertinent professional certification status demographics were also revealed. The majority of participants, 36%, reported to be certified Histology technicians (HT), followed by Pathologists' Assistant (PA), 18%, Histotechnologist 12%, Clinical Laboratory Scientist/Medical Laboratory Scientist 6%, Pathologist 5%, and Cytotechnologist (CT) 2%. The finalized demographic breakdown of participants professional certification status can be seen in Table 4.

Table 4. Professional Certification Status



An observation important to interpreting this study's results is the frequency of respondents who were categorized as being in the in-group (16%) versus those in the out-group (84%). Non-certified AP staff, Non-certified Histology Technician, Certified Histology Technician (HT), Certified Histotechnologist (HTL), Certified Pathologists' Assistant (PA), Non-certified Grossing Technician, Qualification in Immunohistochemistry (QICH) were all merged into the out-group category. Pathologist, and Certified Cytotechnologist (CT), Certified Medical

Laboratory Technician (MLT), Certified Clinical Laboratory Scientist / Medical Laboratory Scientist (CLS/MLS), Certified in limited laboratory specialty (e.g., Chemistry, Hematology, Microbiology) were all merged into the in-group classification, shown in Table 5 below. A dummy variable labeled as group classification was created to identify respondents by group classification.

Table 5. Group Classification

Group Classification	N	%
In-group	43	16%
Out-group	222	84%

3.4 Dependent Variable T-Test

A t-test was performed as part of the data analysis to compare the means of those participants in the in-group and out-group to determine whether there was a statistically significant difference between these means (see Table 6 and 7).

Table 6. Group Statistics

Group Class		N	Mean	Std. Deviation	Std. Error Mean
Institution	In-group	43	2.6744	2.6790	0.4085
	Out-group	222	2.2658	1.8970	0.1273
Role	In-group	43	2.7442	1.1971	0.1826
	Out-group	222	1.5586	0.8634	0.0579
Experience	In-group	43	16.9767	8.9960	1.3719
	Out-group	221	13.7240	10.1655	0.6838
Age	In-group	39	5.4872	1.2112	0.1939
	Out-group	189	5.2646	1.3384	0.0974
Gender	In-group	39	1.6154	0.4929	0.0789
	Out-group	188	1.8191	0.4952	0.0361
Edu	In-group	39	5.8974	0.9402	0.1505
	Out-group	189	4.9524	1.1074	0.0806
Ethnicity	In-group	39	3.4103	2.7117	0.4342
	Out-group	189	2.7354	2.1669	0.1576
PJ	In-group	40	5.6571	0.9981	0.1578
	Out-group	200	4.9321	1.3632	0.0964
Improvement	In-group	39	5.7607	0.8531	0.1366
	Out-group	189	5.7090	0.8879	0.0646
DJ	In-group	40	4.6938	1.6837	0.2662
	Out-group	199	4.3656	1.8443	0.1307
ING_ID	In-group	38	5.9474	0.6846	0.1111
	Out-group	192	5.6632	1.0078	0.0727
Turnover	In-group	39	3.3436	1.5546	0.2489
	Out-group	191	3.5665	1.6247	0.1176
Altruism	In-group	38	5.8421	0.7236	0.1174
	Out-group	194	5.3907	1.1230	0.0806
OUTG_ID	In-group	39	1.5385	0.6606	0.1058
	Out-group	193	1.9655	1.1097	0.0799
Performance	In-group	39	5.3419	1.0713	0.1715
	Out-group	192	4.4089	1.2891	0.0930
JS	In-group	39	4.1218	1.2564	0.2012
	Out-group	197	3.4848	1.2631	0.0900

Table 7. Independent Samples Test

		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interva of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
Institution	Equal variances assumed	6.5609	0.0110	1.2011	263.0000	0.1154	0.2308	0.4087	0.3402	-0.2613	1.078
	Equal variances not assumed			0.9550	50.4638	0.1721	0.3442	0.4087	0.4279	-0.4507	1.268
Role	Equal variances assumed	**17.2238**	0.0000	7.6948	263.0000	**0.0000**	0.0000	**1.1856**	0.1541	0.8822	1.489
	Equal variances not assumed			6.1903	50.7919	0.0000	0.0000	1.1856	0.1915	0.8011	1.570
Experience	Equal variances assumed	**1.1463**	0.2853	1.9540	262.0000	**0.0259**	0.0518	**3.2528**	1.6646	-0.0250	6.530
	Equal variances not assumed			2.1220	64.6996	0.0188	0.0377	3.2528	1.5329	0.1912	6.314
Age	Equal variances assumed	0.1509	0.6980	0.9605	226.0000	0.1689	0.3378	0.2226	0.2318	-0.2341	0.679
	Equal variances not assumed			1.0259	58.8087	0.1546	0.3091	0.2226	0.2170	-0.2116	0.656
Gender	Equal variances assumed	**4.3830**	0.0374	-2.3405	225.0000	**0.0101**	0.0201	**-0.2038**	0.0871	-0.3753	-0.032
	Equal variances not assumed			-2.3477	55.0891	0.0113	0.0225	-0.2038	0.0868	-0.3777	-0.029
Edu	Equal variances assumed	1.5338	0.2168	4.9702	226.0000	**0.0000**	0.0000	**0.9451**	0.1901	0.5704	1.319
	Equal variances not assumed			5.5349	61.8492	0.0000	0.0000	0.9451	0.1707	0.6037	1.286
Ethnicity	Equal variances assumed	**7.7185**	0.0059	1.6919	226.0000	**0.0460**	0.0920	**0.6748**	0.3988	-0.1111	1.460
	Equal variances not assumed			1.4608	48.5036	0.0753	0.1505	0.6748	0.4619	-0.2538	1.603
PJ	Equal variances assumed	**6.6448**	0.0105	3.1944	238.0000	**0.0008**	0.0016	**0.7250**	0.2270	0.2779	1.172
	Equal variances not assumed			3.9207	71.5777	0.0001	0.0002	0.7250	0.1849	0.3563	1.093
Improvement	Equal variances assumed	0.0045	0.9464	0.3332	226.0000	0.3697	0.7393	0.0517	0.1551	-0.2540	0.357
	Equal variances not assumed			0.3421	56.3188	0.3668	0.7336	0.0517	0.1511	-0.2510	0.354
DJ	Equal variances assumed	2.4024	0.1225	1.0413	237.0000	0.1494	0.2988	0.3282	0.3152	-0.2927	0.949
	Equal variances not assumed			1.1065	59.3990	0.1365	0.2730	0.3282	0.2966	-0.2652	0.921
ING_ID	Equal variances assumed	**4.8305**	0.0290	1.6624	228.0000	**0.0489**	0.0978	**0.2842**	0.1709	-0.0526	0.621
	Equal variances not assumed			2.1407	72.9511	0.0178	0.0356	0.2842	0.1327	0.0196	0.548
Turnover	Equal variances assumed	0.0569	0.8117	-0.7863	228.0000	0.2162	0.4325	-0.2229	0.2835	-0.7815	0.335
	Equal variances not assumed			-0.8097	56.2804	0.2108	0.4215	-0.2229	0.2753	-0.7743	0.328
Altruism	Equal variances assumed	**4.2244**	0.0410	2.3806	230.0000	**0.0091**	0.0181	**0.4514**	0.1896	0.0778	0.825
	Equal variances not assumed			3.1697	76.8660	0.0011	0.0022	0.4514	0.1424	0.1678	0.735
OUTG_ID	Equal variances assumed	**4.9205**	0.0275	-2.3189	230.0000	**0.0106**	0.0213	**-0.4270**	0.1841	-0.7898	-0.064
	Equal variances not assumed			-3.2215	88.0332	0.0009	0.0018	-0.4270	0.1325	-0.6904	-0.163
Performance	Equal variances assumed	**3.8892**	0.0498	4.2310	229.0000	**0.0000**	0.0000	**0.9330**	0.2205	0.4985	1.367
	Equal variances not assumed			4.7813	62.5634	0.0000	0.0000	0.9330	0.1951	0.5430	1.323
JS	Equal variances assumed	**0.0255**	0.8732	2.8800	234.0000	**0.0022**	0.0043	**0.6370**	0.2212	0.2013	1.072
	Equal variances not assumed			2.8904	54.3066	0.0028	0.0055	0.6370	0.2204	0.1952	1.078

Statistically, In-group and Out-group populations are significantly different with respect to PJ. Perceived PJ of those in in-group classification is significantly higher than those in out-group with p-value of < .001. The perceived ING-ID for those in the in-group is significantly higher than those in the out-group p-value of 0.029, and perceived OUTG-G is significantly higher in those in the out-group classification than those in the in-group with p-value of 0.028. Service quality and perceived unit performance of those in the in-group classification is significantly higher than those in the out-group with a p-value of 0.05. Job satisfaction of those in the in-group class is higher than those in the out-group with p-value of .002. The means for both in-group and out-group were both low, with the average for those in the in-group being 4.12, and those in the out-group being 3.48 closer to. These results suggest that although there is a general low JS

amongst lab professionals, there is a 0.63702 difference in JS between those in the in-group versus those in the out-group.

3.5 Regression Analyses

Based on the results reported in Tables 6 and 7, the means were significantly different. The next step was to confirmed by running a generalized linear model (GLM) regression techniques to compare in-group and out-group classifications with all-other categories and assess the hypothesized relationship between the constructs. In addition, the constructs in the model were evaluated for multicollinearity, parameter estimates to access construct contributions, statistical significance, and model diagnostics. Table 8 below provides a summary of the hypothesis results with standardized Betas.

Table 8. Summary of Hypothesis Results with Standardized Betas

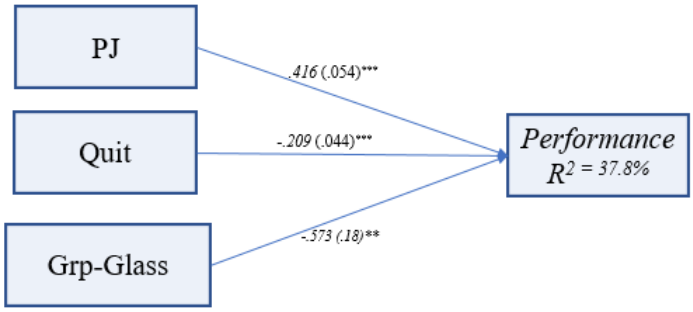
Test	Hypothesis Relationship	Parameter Estimate	t Value	β	Standardized β	Supported
H1a	Grp Class ➔ PJ	-0.72**	3.13	-0.725	-0.203	Yes
H1b	Grp Class ➔ DJ					N/A
H1c	Grp Class ➔ JS	-0.63**	2.88	-0.63	-0.185	Yes
H1d	Grp Class ➔ Professional ID	0.42*	-2.31	0.42	0.151	Yes
H2a	Grp Class ➔ DJ ➔ Perform					N/A
H2b	Grp Class ➔ PJ➔ Perform	0.416***	7.715	0.416	0.431	Yes
H2c	Grp Class ➔ JS➔ Perform					N/A
H2d	Grp Class ➔ Professional ID ➔ Perform					N/A
H3a	Grp Class ➔ DJ ➔ Quitting	-0.107*	-2.012	-0.107	-0.120	Yes
H3b	Grp Class ➔ PJ ➔ Quitting					N/A
H3c	Grp Class ➔ JS ➔ Quitting	-0.418***	-5.394	-0.418	-0.330	Yes
H3d	Grp Class ➔ Professional ID ➔ Quitting	-0.275**	-2.567	-0.275	-0.165	Yes

*p<0.05, **p<0.01, ***p<0.001

During regression analysis, all the demographics became insignificant, so they were dropped, except for race, which was retained in the post-hoc analysis. The findings show that hypothesis H1a and H1c were supported in the data, where participants that were part of the out-group had a lower perception of procedural justice (β= -.725) than those participants in the in-group,

followed by lower job satisfaction (β= -.637). Additionally, professional identification was also influenced by the group classification, as participants that were in the out-group had higher out-group professional identification as compared to those participants in the in-group -.427 units, supporting H1d.

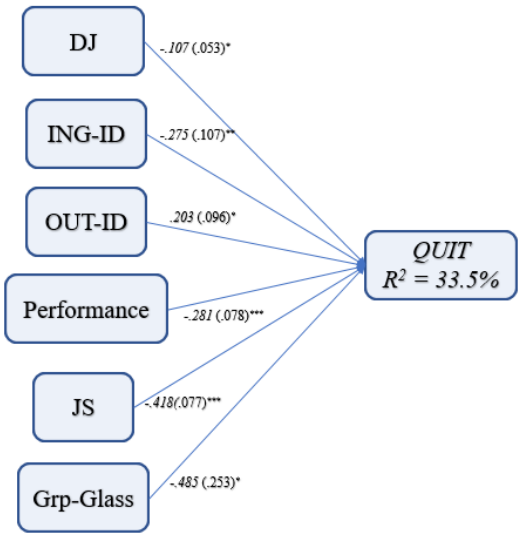
One of our more interesting findings centered around perceptions of disenfranchisement in procedural justice of the out-group and its effects on performance. The analysis supported out-group classification to be associated with decreased performance and higher turnover intention as compared to the in-group. The perceived performance model indicates that out-group professional classification affects perceived performance, so that being in the out-group decreases performance by (β= -.573) supporting H2b. Out-group classification was shown to affect perceived performance through its effects on procedural justice, so that as procedural justice perception increases by 1-unit, performance increases by .416 units, and as turnover intention increases by 1-unit, performance decreases by -.209 units. In addition to the significant path from group classification to performance, we also discovered that both procedural justice and job satisfaction partially medicated the relationship between group classification and performance. In fact, procedural justice perception had an indirect and negative effect (β= -.3575) between group classification and performance. Additionally, job satisfaction was also found to have a negative effect (β= -.1840) in the relationship between group classification and performance. Figure 2 below shows the structural model analysis results.



*p<0.05, **p<0.01, ***p<0.001

Figure 2. Results of the Structural Model Analyses for Performance

The analysis also supported the association between perceptions of disenfranchisement of the out-group with increased employee withdrawal as compared to the in-group. The employee withdrawal model indicates that out-group professional classification affects turnover intention. Turnover intention was significantly predicted by distributive justice, so that as fairness of pay perception increases by 1-unit, turnover intention decreases by -.107 units, there for supporting H3a. Additionally, professional identification also significantly predicted turnover intention, such that as in-group professional identification increases by 1-unit, turnover intention decreases by -.275 units, and as out-group professional identification increases by 1-unit, turnover intention also increases by .203 units, there for supporting H3c. and H3d respectively. Turnover intention was also significantly predicted by performance (β= -.281). Figure 3 below shows the structural model analysis results.



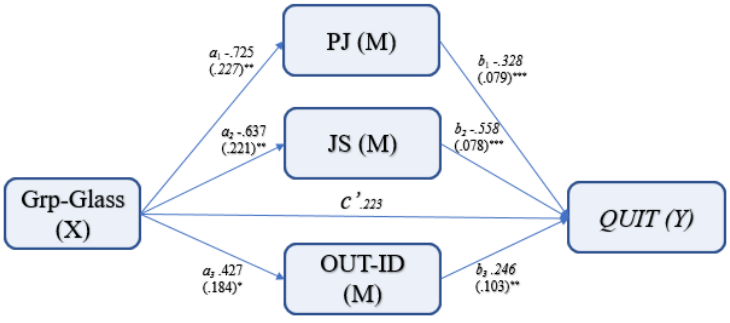
*p<0.05, **p<0.01, ***p<0.001

Figure 3. Results of the Structural Model Analyses for Turnover Intention

Additional analysis discovered the heterogeneity of the population. The outlier data points were associated with one of the demographic groups in the population, in particular Asian Females. After accounting for that demographic group, the group classification becomes highly significant in the employee withdrawal model, so that those in the out-group classification are significantly more likely to have an increase in turnover intention than in-group participants by (β= -1.072). This demographic had no significant effect on the performance model, or on DJ, JS, and PI, however, cultural demographics do have some effect on the perception of procedural justice (PJ) in my sample population. PJ perception was significantly predicted by group class of Caucasians (β= -.706), but not for participants that identified themselves as Asians (β= -.783). This is perhaps due to some difference in the way some individuals of a particular cultural background might shift priorities in respect to their emotions and attitudes as it relates to their perception of procedural fairness. Other cultural demographics effects could not be tested due to the limited data set. Further study into any potential relationship between different cultural demographic

groups and their perception of PJ, DJ, JS and PI could aid managers as they tailor their approach to increase their departments' efficiency and best suit their workforce needs.

In addition to the significant path from group classification to turnover intention, we also discovered that both procedural justice and job satisfaction fully mediated the relationship between group classification and turnover intention. PJ was required to mediate the relationship between group classification and turnover intention. In fact, PJ had a direct and positive effect (β= .2378) between group classification and turnover intention. Additionally, job satisfaction was also found to have a positive effect (β= .3554) in the relationship between group classification and turnover intention. Although out-group identification was significantly predicted by group class (β = .427), and turnover intention was predicted by out-group identification (β = .246), the analysis indicated that there was no significant mediation effect between group class, professional identification, and turnover intention. Even though there was no significant mediation effect identified within this data set, there is still a significant direct effect in the interaction between group classification, out-group identification, on turnover intention. In fact, the analysis shows that participants that were classified as being part of the out-group had significantly lower levels of professional identification and significantly higher levels of turnover intention. Figure 7 below shows the Sobel test analysis results for turnover intention.



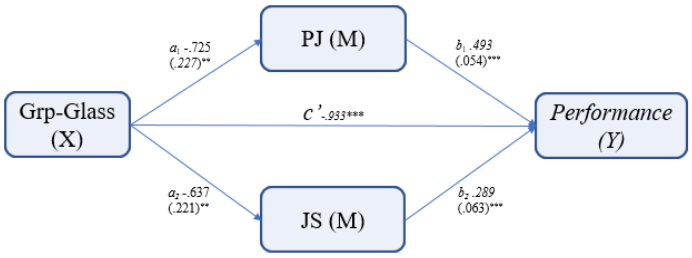
Indirect effect of X on Y through $M = a\ b$ $a_1 * b_{1=.2378**}$

Direct effect of X on $Y = c'$

$a_2 * b_2$=.3554**

Figure 4. Mediation for Turnover Intention

The mediation effect of both PJ and JS on group classification to performance shows a partial because there still is a significant path from group class to performance. Figure 10 below shows the Sobel test analysis results for performance.



Indirect effect of X on Y through $M = a\ b$ Direct effect of X on $Y = c'$

$a_1 * b_1$= -.3574.**

$a_2 * b_2$= -.1840**

Figure 5. Mediation for Performance

4. CONCLUSIONS AND RECOMMENDATIONS

4.1 Summary of Results

The results of this study are instructive in highlighting healthcare industry challenges and the role those governmental regulations play in influencing employee's environment and their behavior. CLIA 88 created an environment of categorization of "in-group" and "out-group." The findings in this study support that an environment of disenfranchisement of the out-group is a significant contributing factor towards the outlined errors in pathology diagnostic services, due to lower employee performance and quality of services, and higher turnover intentions.

4.2 Implications for Organizations

There have been significant advances in science and medicine that have and will continue to transform laboratory medicine (Curry, et al, 2014). The processing of tissue samples has become highly complex, and with personalized medicine becoming the standard of care, the technical quality of the tissue specimen can directly impact quality patient outcomes. With change comes the need for educational routes that align with other clinical laboratory sciences, including a greater number of accredited degree programs, culminating in an Associates, Baccalaureate, or Master's degree. These programs focus on advanced scientific KSA training, as well as critical thinking and decision making necessary for national certification.

Although industry structures are beyond management's control, leaders can devise internal structures that aim to drive innovation, creativity, and upscale/educate our workforce to compete. Individual and collective behavior changes as their environment changes (Vessey & Ward, 2013). Organizations must educate staff and engage them to build support. It is important for the entire organization to be involved in the process of building and incorporating career development programs into their business processes. Sustaining an increased level of staff engagement through interacted learning action loops may also moderate the negative effects of the regulatory landscape outlined in this paper, decreasing complacency and inertia, and enabling organizations to actively pursue increased levels of alignment. Organizational alignment might be best looked at not as a destination, but a journey to constantly improve its capabilities to survive and also thrive in their environment. Prior research has shown that greater alignment among various organizational components, or business domains, produces superior performance (Chan, Huff, Barclay & Copeland 1997, Sabherwal and Kirs 1994). This would enable senior leadership to oversee and analyze how to best align the core services of the lab to position it to maximize HPs' professional development and drive long-term performance outcomes.

4.3 Contributions to Theory

This paper contributes to the team effectiveness literature by elucidating the processes by which regulatory policy may create a work environment that may yield unintended negative outcomes on employee behavior and motivation. Our model linked social identification theory, as it pertains to how people define their self-identity in part based to some extent with the social group, they are part off (Islam, 2014; Tajfel, 78; Tajfel & Turner, 79), and justice theory as a means to contextualize the perceptions of those social groups. This study brings it all together, and it does so in the context of government regulatory exclusion effects on healthcare industry to provide new insights into social and behavioral sciences by demonstrating how this naturally occurring phenomena has influenced employee behavior and engagement. Aside from the context of our research, perhaps our most notable contribution to the literature stems from our findings around potential sociodemographic differences and its effects on procedural justice, and distributive justice perception withing our data set.

4.4 Limitations and Future Research

We would have preferred to gather organizational performance data from industry, but there are HIPAA and other confidentiality restrictions that makes gathering of lab performance data challenging. With that said, we had to utilize perceived performance and theoretically link that with actual performance. Further research could collect actual performance measures and compare it to the perceived performance.

While the number of observations satisfied the criteria for intricate measurement, we are limited to the number of participants. The final population sample in this study was unevenly distributed due to unplanned oversampling of the out-group population (84%) as compared to the in-group

population (16%). Even with this sampling discrepancy, the GLM statistical analysis yielded a high degree of statistical significance. The results should be further studied, future research should aim to increase the population sample size and target a more even distribution between the groups.

4.5 Conclusions

This study explored the consequences that have resulted from the current regulatory environment of staff working in Anatomic Pathology laboratories. This research shows us the effects that government regulatory exclusion has had on the "discredited" careers in terms of employees' perception of unfair work environment, lower job satisfaction and lower professional identification. It also supports finding through SIT by making connections between these social groups that were manufactured by the current regulatory framework which outlines those professions that were recognized as "accredited" or "in-group" and those who were not recognized by virtue of omission as "discredited" or "out-group." It highlights the pivotal roles that these social groups play in increasing dissatisfaction or resentment and predicting employees' behaviors, negative feelings of justice, lower performance, lower job satisfaction, and high employee withdrawal. Our study merely scratches the surface as to the potential for new and interesting discoveries in this area, in particular as it relates to performance in a healthcare setting as a means to decrease diagnostic errors. By recognizing the role of government regulatory exclusion and its effects, this model can assist organizations in their ability to effectively adopt a knowledge-based approach to derive better decisions and improve performance.

References

Abt, A. B., Abt, L. G., & Olt, G. J. (1995). The effect of interinstitution anatomic pathology consultation on patient care. *Archives of pathology & laboratory medicine, 119*(6), 514-517.

Adams, K., Hean, S., Sturgis, P., & Clark, J.M. (2006), Investigating the factors influencing professional identity of first-year health and social care students. *Learning in Health and Social Care, 5*(2), 55-68. https://doi.org/10.1111/j.1473-6861.2006.00119.x

Adams, J. S. (1965). *Inequity in social exchange*. In L. Berkowitz (Ed.), Advances in experimental social psychology (2nd ed.). New York: Academic Press. (pp. 267-299).

Arora, V., Moriates, C., & Shah, N. (2015). The challenge of understanding health care costs and charges. *AMA journal of ethics, 17*(11), 1046-1052. Black, W. C. (2015). Computed tomography screening for lung cancer in the National Lung Screening Trial: a cost-effectiveness analysis. Journal of thoracic imaging, 30(2), 79.

Bateman, T. S., & Organ, D. W. (1983). Job satisfaction and the good soldier: The relationship between affect and employee "citizenship." *Academy of Management Journal, 26*, 587-595.

Borman, W. C. (1991). *Job behavior, performance, and effectiveness*. In M. D. Dunnette & L. M. Hough (Eds.), Handbook of industrial and organizational psychology, 2, (2), 271-326. Palo Alto, CA: Consulting Psychologists Press.

Bray, F. (2014). Transitions in human development and the global cancer burden. *World Cancer Report,* 34. Editors Bernard W Stewart, Christopher P Wild. Lyon: Naturaprint. https://www.drugsandalcohol.ie/28525/1/World%20Cancer%20Report.pdf

Brehm, J. (1972). *Responses to loss of freedom: A theory of psychological reactance*. New York: General Learning Press.

Brief, A. P. (1998). *Attitudes in and around organizations*. Thousand Oaks, CA: Sage.

Brief, A. P., & Roberson, L. (1989). Job attitude organization: An exploratory study. *Journal of Applied Social Psychology, 19*, 717-72.

Bring Histotechnology Processes Under CLIA's Oversight. New York State Histotechnological Society Updates |. (2018, March 10). Retrieved August 9, 2022, from http://www.nyhisto.com/announcements/

Brott, P. E., & Myers, J. E. (1999). Development of a professional school counselor identity: A grounded theory. *Professional School Counseling, 2*(5), 339-348.

Carsten, J. M., & Spector, P. E. (1987). Unemployment, job satisfaction, and employee turnover: A meta-analytic test of the Muchinsky model. *Journal of applied psychology, 72*(3), 374.

Cammann, C., Fichman, M., Jenkins, D., & Klesh, J. (1979). *The Michigan Organizational Assessment Questionnaire*. Unpublished manuscript, University of Michigan, Ann Arbor.

Cammann, C., Fichman, M., Jenkins, G. D., & Klesh, J. R. (1983). *The Michigan Organizational Assessment Questionnaire*. In S. E. Seashore, E. E. Lawler, P. H. Mirvis, & C. Cammann (Eds.), Assessing organizational change: A guide to methods, measures, and practices (pp. 71-138). New York, NY: Wiley.

Chan, Y. E., Huff, S. L., Barclay, D. W., & Copeland, D. G. (1997). Business strategic orientation, information systems strategic orientation, and strategic alignment. *Information systems research, 8*(2), 125-150.

Chalykoff, J., & Kochan, T. A. (1989). Computer-aided monitoring: Its influence on employee satisfaction and turnover. *Personnel Psychology, 42*, 807-834.

Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes, 86*, 278-321.

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O., & Ng, K. Y. (2001a). Justice at the millennium: A Meta-analytic review of 25 years of Organizational Justice Research. *Journal of Applied Psychology, 86*(3), 425-445. https://doi.org/10.1037/0021-9010.86.3.425

Colquitt, J. A. (2001b). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology, 86*, 386-400.

Cropanzano, R., Byrne, Z., Bobocel, D. R., & Rupp, D. (2001). Moral virtues, fairness heuristics, social entities, and other denizens of organizational justice. *Journal of Vocational Behavior, 58*(2), 164-209.

Curry, J. L., Davies, M. A., Calderone, T. L., Nathanson, K., Prieto, V. G., & Gershenwald, J. E. (2014). Tissue Resources for Clinical Use and Marker Studies in Melanoma. In: Thurin, M., Marincola, F. (eds). *Molecular Diagnostics for Melanoma. Methods in Molecular Biology (Methods and Protocols), 1102*. Humana Press, Totowa, NJ.

Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues, 31*, 137-150.

Dollinger, M. J., & Golden, P. A. (1992). Interorganizational and collective strategies in small firms: Environmental effects and performance. *Journal of Management, 18*, 695-715.

Elliot, A. J., & McGregor, H. A., (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*, 501-514.

Epstein, J. I., Walsh, P. C., & Sanfilippo, F. (1996). Clinical and cost impact of second-opinion pathology: review of prostate biopsies prior to radical prostatectomy. *The American journal of surgical pathology, 20*(7), 851-857.

Farrell, D., & Stamm, C. L. (1988). Meta-analysis of the correlates of employee absence. *Human Relations, 41*, 211–227.

Feldman, D. C. (1996). The nature, antecedents and consequences of underemployment. *Journal of Management, 22*(3), 385-407.

Fernandes, C., & Awamleh, R. (2006). Impact of organizational justice in an expatriate work environment. *Management Research News, 29*(11), 701-12.

Fisher, C. D., & Gitelson, R. (1983). A meta-analysis of the correlates of role conflict and ambiguity. *Journal of Applied Psychology, 68*, 320-333.

Folger, R., & Cropanzano, R. (1998). *Organizational justice and human resource management*. Sage, Thousand Oaks, CA.

Gazzola, N., & Smith, J. D. (2007). Who do we think we are? A survey of counsellors in Canada. *International Journal for the Advancement of Counselling, 29*(2), 97-110. doi:10.1007/s10447-007-9032-y

Giddens, A. (1984). *The constitution of society: introduction of the theory of structuration*. University of California Press.

Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education.

Hair, J. F. (1995). *Multivariate data analysis with readings* (Fourth ed.). Englewood Cliffs, N.J: Prentice Hall.

Hanisch, K. A., & Hulin, C. L. (1991). General attitudes and organizational withdrawal: An evaluation of a causal model. *Journal of Vocational Behavior, 39*, 110-128.

Hogg, M. A., & Williams, K. D. (2000). From I to We: social identity and the collective self. *Group Dynamics: Theory, Research and Practice, 4*(1), 81-97.

Homans, G. (1961). *Social behaviour: Its elementary forms*. London: Routledge & Kegan Paul.

Hurst, J., Nickel, K., & Hilborne, L. H. (1998). Are Physicians' Office Laboratory Results of Comparable Quality to Those Produced in Other Laboratory Settings? *Journal of the American Medical Association, 279*(6), 468-471. doi:10.1001/jama.279.6.468

Iaffaldano, M. T., & Muchinsky, P. M. (1985). Job satisfaction and job performance: A meta-analysis. *Psychological Bulletin, 97*, 251-273.

Islam, G. (2014). Social Identity Theory - ResearchGate. (n.d.). Retrieved October 29, 2021, from https://www.researchgate.net/profile/Gazi-Islam 2/publication/281208338_Social_Identity_Theory/links/55db57ec08ae9d6594935f59/Social-Identity-Theory.pdf.

Jackson, S. E., & Schuler, R. S. (1985). A meta-analysis and conceptual critique of research on role ambiguity and role conflict in work settings. *Organizational Behavior and Human Decision Processes, 36*, 16-78.

Jex, S. M., Beehr, T. A., & Roberts, C. K. (1992). The meaning of occupational stress items to survey respondents. *Journal of Applied Psychology, 77*, 623-628.

John, T. D., & Mark, A. H. (1996). The Impact of Human Resource Management Practices on Perceptions of Organizational Performance. *Academy of Management Journal, 39*(4) 949-969.

Judge, T. A., & Bono, J. E. (2001). Relationships of core self-evaluations traits- self-esteem, generalized self-efficacy, locus of control, and emotional stability-with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology, 86*, 80-92.

Judge, T. A., Thoresen, C. J., Bono, J. E., & Patton, G. K. (2001). The job satisfaction-job performance relationship: A qualitative and quantitative review. *Psychological Bulletin, 127*, 376-407.

Leventhal, G. S. (1976). *The distribution of rewards and resources in groups and organizations*. In L. Berkowitz & W. Walster (Eds.), Advances in experimental social psychology (Vol, 9, pp. 91-131). New York: Academic Press.

Leventhal, G. S. (1980). *What should be done with equity theory? New approaches to the study of fairness in social relationships*. In K. Gergen, M. Greenberg, & R. Willis (Eds.), Social exchange: Advances in theory and research (pp. 27-55). New York: Plenum.

Leventhal, G. S., Karuza, J., & Fry, W. R. (1980). *Beyond fairness: A theory of allocation preferences*. In G. Mikula (Ed.), Justice and social interaction (pp. 167-218). New York: Springer-Verlag.

Lunz, M. E., Castleberry, B. M., James, K., and Stahl, J. (1987). The Impact of the Quality of Laboratory Staff on the Accuracy of Laboratory Results. *Journal of the American Medical Association, 258*(3), 361-363. doi:10.1001/jama.1987.03400030077036

Mariotto, A. B., Enewold, L., Zhao, J., Zeruto, C. A., & Yabroff, K. R. (2020). Medical care costs associated with cancer survivorship in the United States. *Cancer Epidemiology and Prevention Biomarkers, 29*(7), 1304-1312.

Mariotto, A. B., Yabroff, K. R., Shao, Y., Feuer, E. J., & Brown, M. L. (2011). Projections of the Cost of Cancer Care in the United States: 2010-2020. JNCI: *Journal of the National Cancer Institute, 103*(2), 117-128. https://doi.org/10.1093/jnci/djq495

Meier, F. A., Zarbo, R. J., Varney, R. C., Bonsal, M., Schultz, D. S., Vrbin, C. M., ... & Raab, S. S. (2008). Amended reports: development and validation of a taxonomy of defects. *American journal of clinical pathology, 130*(2), 238-246.

Meier, F. A., Varney, R. C., & Zarbo, R. J. (2011). Study of amended reports to evaluate and improve surgical pathology processes. *Advances in anatomic pathology, 18*(5), 406-413.

Moorman, R. H. (1991). Relation between organizational justice and organizational citizenship behaviors: does fairness perception influence employee citizenship? *Journal of Applied Psychology, 76*(6), 845-55.

Nelson, K. W., & Jackson, S. (2003). Professional counselor identity development: A qualitative study of Hispanic student interns. *Counselor Education and Supervision, 43*(1), 2-14.

Niehoff, P. B., & Moorman, H. R. (1993). Justice as a Mediator of the Relationship between Methods of Monitoring and Organizational Citizenship Behavior. *Academy of Management Journal, 36*(3), 527-556.

Occupational regulation - Office of the Legislative Auditor (February 1999) p. xii. https://www.auditor.leg.state.mn.us/ped/pedrep/9905-all.pdf

Omura, M. C., Motamedi, K., UyBico, S., Nelson, S. D., & Seeger, L. L. (2011). Revisiting CT-guided percutaneous core needle biopsy of musculoskeletal lesions: contributors to biopsy success. AJR. *American journal of roentgenology, 197*(2), 457-461.

Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA: Lexington Books.

Organ, D. W., & Near, J. P. (1985). Cognition vs affect in measures of job satisfaction. *International Journal of Psychology, 20*, 241-253.

Paulsen, N., Callan, V. J., Ayoko, O., & Saunders, D. (2013). Transformational leadership and innovation in an R&D organization experiencing major change. *Journal of Organizational Change Management, 26*, 595-610.

Peck, M., Moffat, D., Latham, B., & Badrick, T. (2018). Review of diagnostic error in anatomical pathology and the role and value of second opinions in error prevention. *Journal of clinical pathology, 71*(11), 995-1000.

Petty, M. M., McGee, G. W., & Cavender, J. W. (1984). A meta-analysis of the relationships between individual job satisfaction and individual performance. *Academy of Management Review, 9*, 712-721.

Pisano, G., Bohmer, R. M., and Edmondson, (2001). Organizational differences in rates of learning: Evidence from the adoption of minimally invasive cardiac surgery. *Management Science, 47*, 752-768.

Pratt, M. G., Rockmann, K. W., & Kaufmann, J. B. (2006). Constructing professional identity: The role of work and identity learning cycles in the customization of identity among medical residents. *Academy of Management Journal, 49*(2), 235-262.

Remley, T. P., & Herlihy, B. (2005). *Ethical, legal, and professional issues in counseling* (2nd ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Sabherwal, R., and Kirs. P. (1994). The alignment between organizational critical success factors and information technology capability in academic institutions. *Decision Sciences Journal, 25*(2), 301-330.

Schneider, B., White, S. S., & Paul, M. C. (1998). Linking service climate and customer perceptions of service quality: Tests of a causal model. *Journal of Applied Psychology, 83*, 150-163.

Schoen, L. G. (1989). In search of a professional identity: Counseling psychology in Australia. *The Counseling Psychologist, 17*(2), 332-343. doi: 10.1177/0011000089172011

Smith, C. A., Organ, D. W., & Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of Applied Psychology, 68*, 653-663.

Soni, A. (2015). Statistical brief #470: Trends in the five most costly conditions among the US civilian noninstitutionalized population, 2002 and 2012. *Agency for Healthcare Research and Quality*. Retrieved January 10, 2022, from http://www.meps.ahrq.gov/mepsweb/data_files/publications/st470/stat470.shtml

Spector, P. E., Dwyer, D. J., & Jex, S. M. (1988). Relation of job stressors to affective, health, and performance outcomes: A comparison of multiple data sources. *Journal of Applied Psychology, 73*(1), 11-19. doi:http://dx.doi.org/10.1037/0021-9010.73.1.11

Spector, P. E., & Jex, S. M. (1998). Development of four self-report measures of job stressors and strain: Interpersonal Conflict at Work Scale, Organizational Constraints Scale, Quantitative Workload Inventory, and Physical Symptoms Inventory. *Journal of Occupational Health Psychology, 3*, 356-367.

Strauss, G. (1968). Human relations-1968 style. *Industrial Relations, 7*, 262-276.

Stull, T. M., Hearn, T. L., Hancock, J. S., Handsfield, J. H., & Collins, C. L. (1998). Variation in Proficiency Testing Performance by Testing Site. *Journal of the American Medical Association, 279*(6), 463-467. doi:10.1001/jama.279.6.463

Suliman, A. M. T. (2007). Links between justice, satisfaction and performance in the workplace: A survey in the UAE and Arabic context. *Journal of Management Development, 26*(4), 294-311.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: *a cancer journal for clinicians, 71*(3), 209-249.

Tajfel, H. (1978). *The achievement of inter-group differentiation*. In H. Tajfel (Ed.), Differentiation between social groups (pp. 77-100). London: Academic Press.

Tajfel, H., & Turner, J. C. (1979). *An integrative theory of inter-group conflict*. In W. G. Austin & S. Worchel (Eds.), The social psychology of inter-group relations (pp. 33–47). Monterey, CA: Brooks/Cole.

Tett, R. P., & Meyer, J. P. (1993). Job satisfaction, organizational commitment, turnover intention, and turnover: Path analysis based on meta-analytic findings. *Personnel Psychology, 46*, 259-293.

Thibaut, J., & Walker, L. (1975). *Procedural justice: A psychological analysis*. Hillsdale, NJ: Erlbaum.

Tsung, J. S. (2004). Institutional pathology consultation. *The American journal of surgical pathology, 28*(3), 399-402.

Turner J. C. (1999) *Some current issues in research on social identity and self-categorization theories*. In: Social Identity, Context, Commitment, Content (eds N. Ellemers, R. Spears & B. Doosje), Blackwell Publishers, Oxford.

Vessey. I., and Ward. K. (2013). The Dynamics of Sustainable IS Alignment: The Case for IS adaptivity. *Journal of the Association for Information Systems, 14*, 283-311.

Volmar, K. E., Idowu, M. O., Hunt, J. L., Souers, R. J., Meier, F. A., & Nakhleh, R. E. (2014). Surgical pathology report defects: a College of American Pathologists Q-Probes study of 73 institutions. *Archives of Pathology and Laboratory Medicine, 138*(5), 602-612.

Wheeler, A. R., Gallagher, V. C., Brouer, R. L., & Sablynski, C. J. (2007). When person-organization (mis)fit and (dis)satisfaction lead to turnover: The moderating role of perceived job mobility. *Journal of Managerial Psychology, 22*, 203-219.

World Health Organization. (2015). *Global health observatory data repository*. 2011. Number of deaths (World) by cause. Retrieved January 10, 2022, from http://apps.who.int/gho/data/node.main.CODWORLD?lang=en.

Yabroff, K. R., Mariotto, A., Tangka, F., Zhao, J., Islami, F., Sung, H., Sherman, R. L., Henley, J., Jemal, A., & Ward, E. M. (2021). Annual Report to the Nation on the Status of Cancer, Part 2: Patient Economic Burden Associated with Cancer Care. JNCI: *Journal of the National Cancer Institute, 113*(12), 1670-1682. DOI: 10.1093/jnci/djab192.

Zarbo, R. J., Meier, F. A., & Raab, S. S. (2005). Error detection in anatomic pathology. *Archives of Pathology and Laboratory Medicine, 129*(10), 1237-1245.

ABSTRACT

Telework is not a new phenomenon; however, during the COVID pandemic that began in 2020, individuals and organizations were forced to adopt remote work for the first time in large numbers. Most were able to adapt with varying degrees of success. The experience led some to reexamine workplace policies, practices, and processes. Individuals began to reassess professional expectations, personal needs, the quality of their commutes to the office, and the long-term feasibility and appeal of telework. Evidence suggests that the appeal of telework is negatively related to commuting satisfaction with moderating and contextual factors. Individuals may tolerate long commutes to access a certain quality of life and affordable standard of living. Long commutes may provide access to economic opportunities and professional networking and advancement possibilities. However, the costs of long commutes can be steep in terms of time lost, direct costs of traveling, and various opportunity costs. It can mean less time spent with friends and family, personal interests, and self-care. Commutes can be stressful, arduous, and frustrating and lead to negative mental and physical health outcomes. This paper proposes a multi-dimensional model that identifies the factors that influence commuting satisfaction that in turn influences the appeal of telework. The factors presented that determine commuting satisfaction are commuting enjoyment, strain, safety, impedance, reliability, mode flexibility, and productivity while in transit. The model presented in this research applies the theory of planned behavior as a theoretical basis.

Keywords: Telework, Telecommuting, Commuting

THE RELATIONSHIP BETWEEN THE MULTIPLE DIMENSIONS OF COMMUTING SATISFACTION AND THE APPEAL OF TELEWORKING

INTRODUCTION

Evolution of Telework

The terms *telework* and *telecommute* are often used interchangeably. For the purposes of this research, we use the term *telework*. Its commonly understood meaning is broad enough to encompass different forms of remote work situations. Telework can entail working from home, satellite offices, client offices, co-working spaces, and public spaces such as cafés and libraries. It can be fully remote or hybrid, the latter of which is more common as of 2022 (Wigert and Agrawal 2022). The hybrid mode entails employees spending time both in the office and working from remote locations, usually the home. Hybrid modes of work are the middle ground that may afford employers and employees the benefits of both in-office and remote work modalities while mitigating the downsides of choosing either modality alone.

Notwithstanding the current interest in teleworking, it is not a new phenomenon. The contemporary concept of telework can be attributed to Jack Nilles as he sat in Los Angeles traffic congestion in 1973 thinking about the inefficiencies of commuting to work (Nilles 1976; Nilles 2018). He developed and promoted the idea that teleworking, or rather telecommuting, a term he coined, could offset traffic congestion and conserve resources. His work significantly influenced the telework workplace model.

Since the 1970s, and up until the start of the pandemic, the growth of telework had been relatively modest. Between 1980 and 2014, the percentage of workforce telework

participation grew from approximately 2.25% to just under 4.5% (Ingraham 2016). In addition, only certain industries and job categories, mostly related to well-compensated professional services, had access to telework options (Desilver 2020). The feasibility of telework had increased substantially in recent years, but many had not seemed to notice until the pandemic forced large numbers of workers to isolate and work from home.

The gradual adoption of telework may have been attributable to perceptions that it was inconsistent with workplace norms and processes. Managers may have resisted telework because of a perceived need to "see" their people at work getting things done. It could have reflected a lack of trust in employees. More generally, some organizations may not have discerned sufficient advantages of teleworking, especially without many observable models of best practices to help guide the design of remote work policies. Experimentation can be costly and time consuming, and inertia can be difficult to overcome. Organizations may have seen little reason to upend business practices without compelling reasons and employees may not have considered telework a realistic possibility. The pandemic changed many attitudes. But there are impediments to telework, and it may not be appropriate for every role.

Challenges of Telework

Work-related impediments. There are work-related impediments. Some employees are tethered to equipment, resources, and assets that are more easily managed in centralized locations. Work sometimes requires a great deal of face-to-face interaction to achieve timely responses to questions and to work through ambiguous challenges and task requirements. Individuals may perceive physical visibility as necessary for promotion and advancement. The concern is that out of sight leads to out of mind. Whether the nature of the business is highly

transactional or entails relationship building could influence the perceived appropriateness of telework. Some managers may believe that telework is not suitable for a dynamic task environment characterized by fast paced, non-routine, or high stakes decision making.

Challenge of managing for optimal outcomes. In a remote work environment, managing for optimal outcomes can be difficult and may require new approaches. Processes for employee orientation, training, development, evaluation, and mentoring may need to be rethought. It could be more challenging to integrate teleworkers into teams and foster high levels of collaboration. Roles, responsibilities, and deliverables may require a greater degree clarity and explanation. Many believe that innovation emerges from the serendipitous encounters, casual mingling, and the cross pollination of ideas when employees are co-located. This could be challenging in remote work environments. Indeed, there are some indications that innovation has declined in this era of remote work, but it is not clear if this is the result of telework or the distractions caused by the pandemic (Wingard 2022).

Challenge of creating and sustaining organizational culture. Telework may not be ideal for different types of organizational cultures such as apprenticeship cultures (Hoang et al. 2008). Organizational culture is often learned and strengthened by developing and sustaining relationships and that can be more difficult in remote work environments. A high level of organizational trust is necessary. In remote work environments, it could be more challenging to strengthen the organization's culture and communicate values in ways that normally rely on the observation of patterns of interactions, the architectural and interior design of workspaces, and other forms of in-person signaling and reinforcement.

Challenge of managing boundaries and confidentiality. When working from home, it can be difficult for individuals to manage boundary issues and maintain cognitive, emotional, and behavioral focus. Managing cybersecurity, ensuring privacy, and maintaining confidentiality can be challenging. Company secrets, customer data, and other types of information can be at greater risk outside the controlled environment of a centralized workplace.

Challenge of managing technology. A broader acceptance of telework may have been constrained by the perceived capability of information and communication technologies (ICTs). By the time the pandemic struck in March of 2020, ICTs had substantially improved in terms of usefulness, ease of use, and quality. Broadband availability, project management software, conferencing tools, and other essential ICTs to manage remote work had improved. ICTs had become easier to use, accessible, and more affordable for the average user. Zoom turned out to be a "killer app" during COVID, but there are many other effective tools such as Microsoft Teams. New tools continue to emerge and evolve.

Benefits of Telework

Telework benefits for organizations. The challenges of teleworking should be weighed against the many benefits. Telework enables a distributed workforce that can support greater resiliency during unexpected events and emergencies as was the case during the COVID pandemic (Budnitz et al. 2021; Carvalho et al. 2022; Colbert 2012). Organizations can benefit from reduced overhead costs, especially those related to real estate, including lease, tax, security, maintenance, and cleaning expenses. Flexible workplace telework policies can be a competitive bargaining chip that allows organizations to attract talent without some of the usual geographic constraints. Flexible telework may lead to improved employee retention and reduced

absenteeism. It could lead to a more inclusive workplace and increase employee job satisfaction and engagement.

Benefits of telework for society. Then there are the broader societal benefits. Most commuters in the United States drive themselves to work. Therefore, a greater adoption of telework could lead to reduced traffic congestion and reduced use of fossil fuels, contributing to environmental sustainability goals. Teleworking could mean a willingness to live greater distances away from a centralized workplace that could bring new economic vitality to suburbs, exurbs, and secondary cities and towns by dispersing economic activity away from the center of gravity in large metropolitan business districts (Elle 2020).

Benefits of telework for individual productivity. On a more granular level, teleworking benefits individuals in various ways. It's an opportunity to develop and improve certain communication skills such as writing. It requires developing and improving virtual collaboration skills and learning about new ICTs. Working remotely requires and helps develop self-organization skills. Workers can avoid office distractions, politics, unwanted socializing, and other unproductive dynamics that could be stressful and enervating. This allows for greater focus and concentration and for longer periods of time. Telework can permit greater flexibility and independence to manage schedules, work processes, and downtime for peak performance. An interesting observation during the pandemic was that many reported high productivity while working from home (Wingard 2022). Time away from the office environment may stimulate creative thinking or afford time to recharge. Telework can lead to an overall sense of wellbeing.

Benefits of telework for inclusiveness. One could argue that modern office spaces with open layouts are really designed for extroverts. Therefore, teleworking could be a better fit for introverts. In addition, telework could make the workplace more inclusive for certain marginalized groups. Women, minorities, and certain age groups could feel more at ease collaborating virtually. Behind a Zoom screen certain differences may not be as salient. Telework policies can make work more inclusive for those with chronic health, mobility, or disability issues. It can reduce the need for taking time off because of illness, caretaking responsibilities, and weather events. In general, teleworking can afford workers the flexibility to better manage their personal lives.

The New Normal

Municipal and business leaders want to reanimate central business districts after the pandemic shutdowns. There is concern about the lack of business activity, lost tax revenue, and transit revenue associated with pre-pandemic patterns of work. Many businesses have substantial real estate exposure and empty office spaces can hit the bottom line hard. As of May 2022 the majority of workers in New York City, which contains the largest business district in the United States, had yet to return to the office full time (Goldberg 2022). This despite the efforts of local governments (Barron 2022) and companies (Cerullo 2023) to coax people, sometimes aggressively, back into the office.

Recent reports suggest that the level of remote work has stabilized and may not change (Barrero et al. 2023). This is forcing many firms to grapple with the idea of fully remote and hybrid work modes as a permanent feature of the workplace. During the pandemic large numbers of employees realized the benefits of teleworking for the first time and are resisting

going back to the previous status quo. It is important to understand individual motivations to telework. It seems that not having to commute is a major factor.

Quality of the Commute

The quality of a commute, taking into account moderating factors, can have an effect on mental and physical wellbeing (Chng et al. 2016; Clark et al. 2020; Henke et al. 2016; Jun et al. 2019; Kahneman et al. 2004; Stone and Schneider 2016). It can affect sleep, diet, stress levels, depression, and overall life satisfaction. The nature and quality of commuting can also have an impact of commuting on productivity (Ma and Ye 2019).

Commuting results in a loss of time and money. Those are certainly factors driving the appeal of telework, but we believe it is more than that. It is about the overall quality of the commute and there are several factors to consider, which we explore further below. In addition, the appeal of and intention to telework will be influenced by normative expectations and having the choice to telework.

In the theoretical development section, we develop a multidimensional model for commuting satisfaction. We theorize the relationships. The variables are integrated into a model based on the theory of planned behavior, which can be used to explain behavioral intentions. We discuss the implications of the proposed model. We provide some concluding remarks.

THEORETICAL DEVELOPMENT

Theory of Planned Behavior overview

The theory of planned behavior can be used to understand whether an individual will decide to telework (Ajzen 1991). According to the theory of planned behavior, intention precedes behavior. The stronger the intention to telework, the more likely an individual will do so. In turn, intentions are influenced by subjective norms, behavioral control, and attitudes.

Telework Perceived Behavioral Control

Perceived behavioral control is the individual's perceived ability to perform a given behavior. The focus here is on teleworking behavior. Perceived behavioral control is a function of self-efficacy and the external factors that can lead one to believe they can telework. An individual may find the idea of teleworking appealing but may not have the skills, tools, or the supportive environment. Telework requires a degree of expertise with the right mix of ICTs. Telework requires the ability to manage certain types of work processes in a remote work environment. It requires an ability to manage boundaries with respect to time, space, work, and private life. Telework requires flexible and supportive workplace policies. It often requires the ability to self-manage and work independently.

Proposition 1. Behavioral control with respect to telework is positively related to telework intention.

Telework Subjective Norm [OKAY]

Subjective norm is an individual's perception about the particular behavior, which is influenced by the judgment of significant others (Ajzen 1991). Subjective norm with respect to telework is

the degree to which people important to the individual believes that a particular modality is acceptable and appropriate. Subjective norms are determined by the normative beliefs concerning the expectations and behaviors of significant others. The significant others could be colleagues, professional peers, supervisors, teachers, friends, and relatives. Depending on the individual and the circumstances, the influence of specific significant others will vary.

An individual who is enthusiastic about teleworking may be discouraged from pursuing such a work arrangement if it is not deemed appropriate. An individual can assess the social norms surrounding telework through interactions with colleagues and professional peers. They can judge subjective norms from the organization's culture or by observing industry norms.

Proposition 2. Subjective norms regarding telework will be positively related to telework intention.

Telework Attitude

Telework attitude is the degree to which an individual has a favorable evaluation of telework. In other words, it is the degree to which an individual finds teleworking appealing. The theory of planned behavior suggests that an individual's attitude toward a behavior is determined by his or her accessible beliefs about the behavior. If an individual has positive expectations of teleworking, s/he will be inclined to adopt the behavior.

We discussed many of the benefits and challenges of teleworking. A net benefit should encourage positive attitudes toward teleworking. There are indications that attitude toward teleworking could also be influenced by individual attributes such as an individual's age (Raišienė et al. 2021) and gender (Raišienė et al. 2021; Teo et al. 1999). Some women may appreciate the flexibility of telework with childcare in mind. Some men may be more concerned

about a detrimental impact on personal career development for not being in the office. Members of minority groups may feel less at ease and included in the workplace and therefore have a greater preference for telework (Williams et al. 2022). Personality traits may be associated with certain attitudes regarding telework. There is some indication that introverts are less prone to burnout in remote work environments (Meymandpour and Bagheri 2017).

Technology factors are likely to influence attitudes about teleworking. General attributes may include ease of use, usefulness, and enjoyment. Other attributes may be related to system, information, and service quality. Specific technology features that could influence attitudes about the utility of telework include the availability of project management, collaboration tools, storage speed and capacity, cloud-based software and services, video, video conferencing, direct and group messaging, business applications and productivity tools, and calendars.

Proposition 3. Telework attitude (appeal) is positively related to telework intention.

Commuting satisfaction

Commuting satisfaction is the degree to which an individual finds their commute pleasurable. It is a psychological state characterized by feelings of approval, contentment, and well-being.

When workers are asked about the benefits of telework, not having to commute often comes up (Lykiardopoulou 2021; Thompson et al. 2021). This suggests that commuting satisfaction could be a significant factor that influences the appeal of telework.

There are several positive outcomes associated with commuting satisfaction (Gimenez-Nadal and Molina 2019; Kahneman et al. 2004; Liu et al. 2022; Stutzer and Frey 2008). They include less stress, depression, and anxiety. It includes improved wellbeing and life satisfaction.

Commuting satisfaction is often associated with increased job satisfaction, commitment, and engagement.

Commuting satisfaction could be influenced by the length of the commute and its cost. For many, living close to centers of employment such central business districts, is highly desirable, but often unattainable because of affordability issues. That is often the case for middle income workers in many large metropolitan areas such as New York and Los Angeles. As a result, individuals must contend with long, sometimes onerous commutes.

Proposition 4. Commuting satisfaction is negatively related to telework attitude.

Determinants of Commuting Satisfaction

Several factors may influence commuting satisfaction:

- *Commuting enjoyment* (Chatterjee et al. 2020; Ma and Ye 2019; Stone and Schneider 2016)
- *Commuting cost* (Lyons and Chatterjee 2008)
- *Commuting distance* (Chatterjee et al. 2020; Gimenez-Nadal and Molina 2019; Higgins et al. 2018; Liu et al. 2022; Lykiardopoulou 2021; Ma and Ye 2019; Nicholls et al. 2018; Simón et al. 2020; Sposato et al. 2012; Stone and Schneider 2016; Stutzer and Frey 2008; Thompson et al. 2021)
- *Commuting strain* (Kluger 1998; Koslowsky and Aizer 1996; Lucas and Heady 2002; Zhou et al. 2017)
- *Commuting safety* (Allen et al. 2019; Kaplan et al. 2022)
- *Commuting impedance* (Chng et al. 2016; Higgins et al. 2018; Kluger 1998; Koslowsky and Aizer 1996; Novaco et al. 1990; Schaeffer et al. 1988; Sposato et al. 2012)
- *Commuting reliability* (Sposato et al. 2012)
- *Commuting mode flexibility and choice* (Ma and Ye 2019; Schaeffer et al. 1988)
- *Commuting personal productivity*
- *Commuting work productivity*

Commuting Enjoyment

Commuting enjoyment is the degree to which an individual believes that traveling between the home and workplace is an enjoyable experience on its own apart from any practical

considerations. Commuting enjoyment refers to the hedonic aspect of a trip. It is the degree that the trip is pleasurable and fun.

The travel experience can vary. How an individual experiences the trips over time can change. Big city rush hour commutes can be arduous and intense in crowded, standing room only buses and underground metro trains. Carrying heavy backpacks can be exhausting. Driving in heavy city traffic can be enervating. Enjoyment can be impacted by environmental stressors such as noise and crowding. (Park and Evans 2016) On the other hand, an above ground commuter trip in a comfortable, uncrowded, air-conditioned train with comfortable seats and nice scenery can be very relaxing and enjoyable. Enjoyment may increase when traveling with company (Chatterjee et al. 2020). It is an opportunity to socialize and connect.

Proposition 5. Commuting enjoyment perception is positively related to commuting satisfaction.

Commuting Cost

Commuting cost is the individual's perception about the relative costs of commuting between the home and workplace. It includes the direct costs of travel such as bus and train fares. It includes the cost of fuel and tolls. Drivers must take into account vehicle maintenance, insurance, parking fees, depreciation, interest on loans, leases, registration costs, and the occasional traffic summons. Traveling to the office requires budgeting for an appropriate wardrobe, its maintenance, and travel gear. It may also include the cost of prepared foods for breakfast and lunch. Some are more sensitive about costs, either because it represents a high percentage of the home budget or because of psychological anchors.

Proposition 6. Commuting cost perception is negatively related to commuting satisfaction.

Commuting Distance

Commuting distance is the perceived distance between the home and workplace. Distance is best thought of in terms of time and not physical distance. In suburban or rural areas, it is typically much easier to travel longer distances within the same amount of time compared to traveling in more densely populated urban areas. The geography of a commute may also influence distance perceptions as well as the mode of travel. A clean, truck-free parkway surrounded by natural beauty could be a relatively relaxing experience compared to driving along a major, multi-lane highway with commercial traffic. Perceptions of distance could differ.

Long commutes influence happiness and overall subjective wellbeing (Liu et al. 2022). Researchers find that satisfaction decreases with the duration of commute, regardless of the mode used. (Chatterjee et al. 2020). Loss of time is one of the most common complaints of workers and a reason many would choose to telework in addition to fatigue and stress. Lengthier commutes may subject travelers to greater probability of disruptions such as road congestion which are psychologically taxing. Longer commutes can circumscribe travel options. Living close to work may afford the commuter the option to walk, bike, or drive.

Shorter commutes can enhance work-life balance. It leaves more time for family, friends, school, recreation, hobbies, and other activities that may bring enjoyment and fulfillment. Shorter commutes can lead to more opportunities for professional networking. Time saved can be used for self-care such as exercise to manage stress and wellbeing. It leaves more time to recharge mentally, physically, and emotionally. Extensive commuting may lead to more unhealthy lifestyle choices such as drinks after work with colleagues and more prepared

meals that may be less nutritious. Prepared meals are convenient but generally more expensive than home cooked meals.

Note that not all commutes, even lengthy ones, are necessarily bad. They can be therapeutic if it affords an individual time away from a distressing neighborhood. (Jacobs and Brailsford 2019). If that's the case, longer commutes could lead to improved mental health outcomes.

Proposition 7. Commuting distance perception is negatively related to commuting satisfaction.

Perceived Commuting Strain

Commuting strain is the degree to which an individual perceives traveling between the home and workplace is an arduous experience. This construct is about the hassle of commuting.

Some commutes are physically demanding. They can be more difficult, uncomfortable, and exhausting than others. It may entail carrying heavy backpacks, walking, biking, standing, or driving, all of which requires effort. Strain can be caused by bad weather events, crowding, lack of cleanliness, and noise. It can be psychologically demanding because of proximity to difficult and unruly passengers leading to frustration, anger, anxiety, resentment, or stress.

Commutes can be cognitively demanding. It requires planning that may include checking the weather, deciding what to wear, checking on potential delays and assessing alternate travel options, packing for the day, and preparing meals. Commuters may need to quickly adapt to route and scheduling changes in real time, while in transit.

Proposition 8. Commuting strain perception is negatively related to commuting satisfaction

Perceived Commuting Safety

Commuting safety is the degree to which an individual perceives their personal safety is free of risk when traveling between the home and workplace. Concerns about personal safety can be related to crime, health, or the risk of accidents.

During the pandemic, many cities around the country have experienced an uptick in crime. The reasons are still being debated. It could be related to less effective policing, or it could be the result of social stresses. In many cities such as New York, crime has spiked on transit systems and there have been some high-profile incidents. This may have discouraged ridership, which is still well below pre-pandemic levels (Bosman et al. 2022). As the COVID pandemic subsided, new health concerns emerged to the forefront including polio, monkeypox, as well as seasonal flu (Freedman 2022; Johnson 2022).

Proposition 9. Commuting safety perception is positively related to commuting satisfaction.

Perceived Commuting Impedance

Commuting impedance is the degree to which an individual experiences travel impediments such as congestion, transfers, confusing announcements, and lengthy or difficult payment processes when traveling between the home and workplace. If commuting is necessary, a smooth, steady, one seat ride between points is ideal. Anything short of that contributes to impedance, i.e., friction.

Impedance can emerge in different ways. Navigating complex transit systems, inconvenient connections, or having to climb staircases contribute to impedance. Dealing with multiple connections, different payment systems, and multiple or inconvenient transit modes and systems create impedance. Difficult to understand track announcements and schedules

create travel friction. A time-consuming payment process, or a system that relies on physical tickets instead of mobile tokens and payment also creates friction. Lack of express lanes, going through toll plazas, congestion, sharing the road with slow moving vehicles creates impedance. Impedance that reflects crowding and congestion results in stress (Chatterjee et al. 2020)

Impedance has consequences. Long waits and missed departures due to incorrect real-time information can lead commuters to abandon public transit (McMahon 2013). “While improvements in travel time matter for increasing commute satisfaction, it is reductions in travel in congested conditions that matter most, particularly among those susceptible to congestion-related stressors” (Higgins et al. 2018).

Proposition 10. Commuting impedance perception is negatively related to commuting satisfaction.

Perceived Commuting Reliability

Commuting reliability is the degree to which an individual perceives that travel time between the home and workplace is consistent. Unreliable transportation induces stress and forces commuters to include extra slack time in their travel plans.

There are several factors that influence perceived reliability. The frequency awareness of accidents, mechanical breakdowns and delays due to passengers needing special assistance influences perceived reliability. The frequency of unanticipated weather events, construction, closures, emergency repair work can raise system reliability and resiliency concerns. The frequency of disruptions in other transit systems and modes can lead to spillover effects that delays everyone’s commute. Frequent transit announcements about changes may influence reliability perceptions. Since the end of the pandemic, there have been an increase in delays

caused by transportation staffing shortages (Biron 2022). Unpredictability can induce stress (Chatterjee et al. 2020), which could increase the appeal of telework.

Proposition 11. Commuting reliability perception is positively related to commuting satisfaction.

Commuting Mode Flexibility

Commuting mode flexibility is the degree to which an individual perceives they have different transportation options for traveling between the home and workplace. Having options matter for a variety of reasons. It is a way of dealing with unanticipated disruptions and delays. It can provide peace of mind knowing that a disruption does not risk missing work or getting home late. Having transportation options can alleviate the boredom that may arise from an unchanging daily routine.

Having the option of choosing between active and passive travel modes can influence commuting satisfaction (Page and Nilsson 2017). Mode flexibility that allows the possibility of a more active commute can improve the physical health of employees and job performance (Ma and Ye 2019; Page and Nilsson 2017).

Proposition 12. Commuting mode flexibility perception is positively related to commuting satisfaction.

Commuting Personal Productivity

Commuting personal productivity is the degree to which an individual perceives that personal tasks can be managed and completed while traveling between the home and workplace. Making productive use of commuting time can ease the burden of travel. The extent to which an

individual can organize their lives, make plans, check in on friends and family, work on projects, read books, or listen to podcasts is a good use of time that could otherwise be lost to traveling.

Not all forms of commuting afford the opportunities to be productive. It is not possible to read a book while driving. It may be uncomfortable to stand and read in a crowded train. There may not be enough privacy to have certain conversations on the phone. There may be unreliable WIFI access.

Proposition 13. Commuting personal productivity perception is positively related to commuting satisfaction.

Perceived Commuting Job Productivity

C*ommuting job productivity is the degree to which an individual perceives they can manage and accomplish job-related tasks while traveling between the home and workplace.* That can entail getting ready for work, planning the day, and thinking through job-related issues. It may be possible to hold phone-meetings while commuting.

The same set of issues apply as in the case for commuter personal productivity. For example, commuters need WIFI access and a comfortable work environment. They may also need privacy, perhaps even more so for work-related tasks.

Proposition 14. Commuting job productivity perception is positively related to commuting satisfaction.

The complete model is presented below in Figure 1.

MODEL ILLUSTRATION

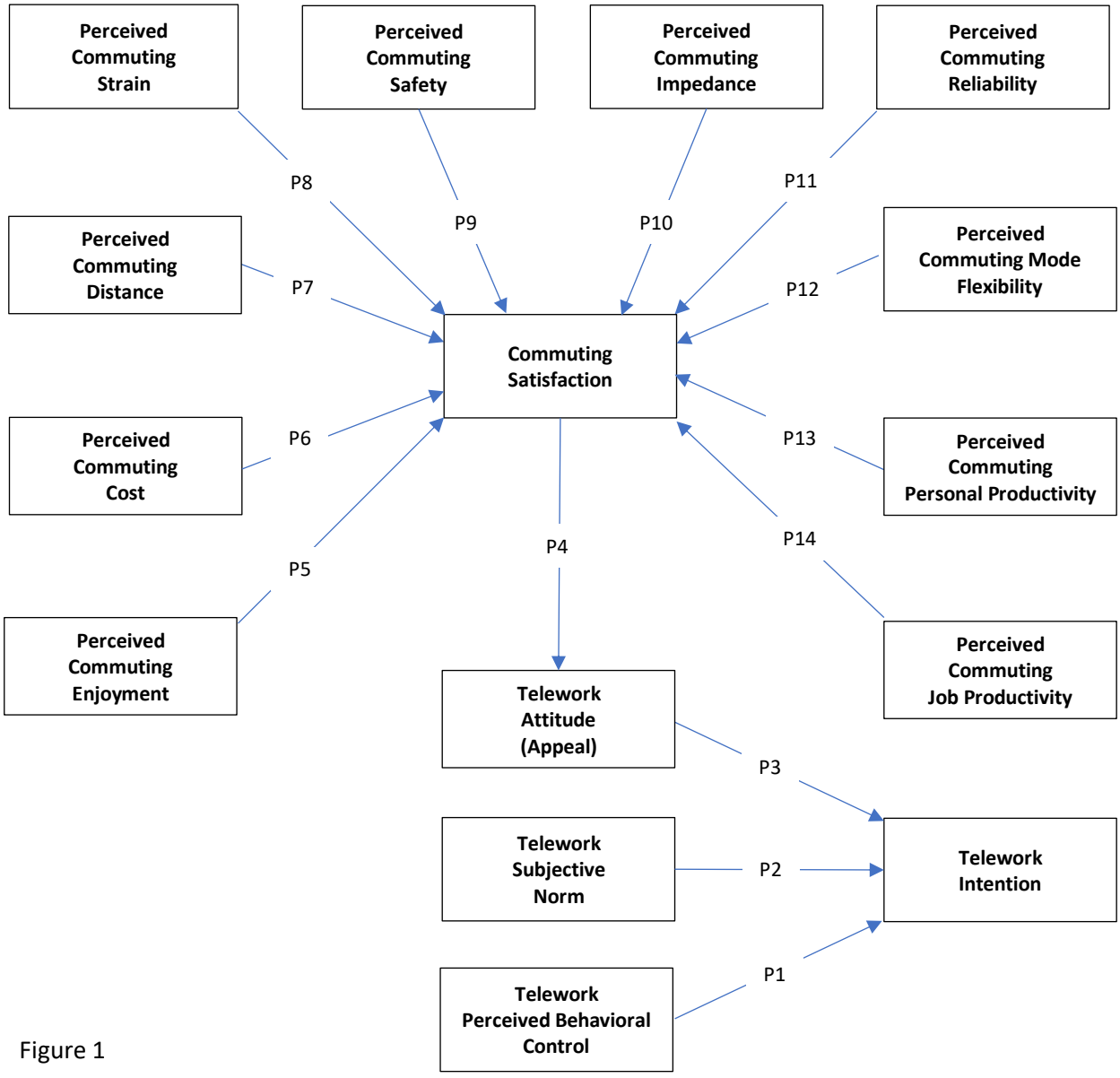


Figure 1

DISCUSSION

The purpose of the proposed model is to better understand all the influential determinants of commuting satisfaction and how they ultimately influence telework attitudes and intentions. It could serve as a guide for future research and policy making. Much of the literature considers how commuting distance matters to individual wellbeing and job satisfaction. Surveys have indicated that commuting is among the least favorite activity for many. Much of the dissatisfaction is a function of the loss of time that could be spent on more meaningful activities. Moreover, the cost of commutes can also take a big bite out of an individual's budget. However, time spent commuting and its cost are not the only important considerations. The issue is about the overall quality of the commute.

An enjoyable commute can offset the burden of a long or expensive commutes and there are many ways in which the pleasure of a commute could be enhanced. It can start with the basics of providing comfortable seating and temperature-controlled environments. It could be achieved by preventing overcrowding, keeping noise to a minimum, providing soothing background music, and maintaining clean, smoke free environments. Special attention can be given to the esthetics of the environment and designing functional spaces with natural lighting. Enjoyable commute can result by providing access to natural, esthetically pleasing scenery, and access to in-transit entertainment. Active travel modes such as walking or biking along protected paths might be more enjoyable than sitting in a car seat fighting rush-hour traffic.

A commute can be arduous even if aspects of the commute are enjoyable. For example, riding a bike is fun for many people, but it can be a demanding experience especially in hilly terrain that requires stamina and ability. Walking to work can be great but only a reasonably

nice day. Inclement weather or having to carrying heavy bags can be a miserable experience. Commuting strain can also be psychological. Driving can be fun but navigating route changes to avoid congestion can be tiring. Commuting strain can be minimized in different ways depending on the modality. E-bikes, moving walkways and escalators, and comfortable seats can ease commuting strain.

One of the challenges of getting people back on public transportation and back to the office is the perceived lack of commuting safety in some cities. Elevated crime remains an issue. But there are different aspects to safety that could be explored. We alluded to concerns regarding COVID, which appears to have declined. But that could be temporary and remain an issue over the long term along with other health and safety concerns. Potential safety issues could be with respect to navigating difficult road conditions. Safety concerns are challenging and expensive to address but are important. The degree to which commuter safety concerns are assuaged could influence the appeal of telework.

Not every commute is smooth. The extent to which commuting friction can be eliminated may influence commuting satisfaction. There are many possibilities for eliminating friction. Mobile payments and trip management apps could speed up a commute. Real time information, advisories, and trip suggestions could help commuters make better decisions to expedite their journeys. Clear signage and announcements can help commuters avoid confusion especially when there are disruptions.

A potential source of frustration for many commuters is variability. Uncertainty regarding travel times create stress and forces commuters to allot a greater amount of time for commuting. Many public transit systems in North America are old or at capacity. Therefore, it

often doesn't require much of a disruption to create serious reverberations that create stressful, time-consuming delays with spillover effects. Avoiding deferred maintenance of transit systems and continual improvement is essential.

Longer commutes may eliminate options for getting to work. When possible, an effort to create alternatives may be a worthwhile investment that could have beneficial effects. Having mode options can reduce the impact of travel disruptions. Alternative ways of getting to work can afford a respite from a dull routine that could be energizing and stimulate creativity. Municipalities are increasingly adding alternatives such as bike lanes to give commuters options, get cars off the road, and to meet sustainability goals. The extent to which commuters are bored with their commutes may influence attitudes toward telework.

The extent to which users can multi-task and be productive while commuting could alleviate some of the burdens of regular travel to the workplace. A way in which productivity can be enhanced on a commute is by ensuring wireless network access all along the commute, especially for the users of public transportation. This would allow commuters to manage and perform work and home related tasks. Even for those who drive, WIKI access would facilitate access to podcasts and other forms of audio consumption. The extent to which commuters can be productive can reduce the sense of time being wasted while traveling between home and work. WIFI and wireless access have greatly improved over the years, but it is not always reliable everywhere, especially when signals are obstructed as is often the case when traveling underground.

CONCLUSION

This is a theoretical paper. Many of the constructs were based on a review of the literature, observations, and recent reporting. The goal was to better understand the motivations for telework. The model presented must be tested to determine efficacy and the extent that it can be used to better under the motivations for telework.

Society seems to have reached an inflection point in which different modalities for telework have become a new normal. After a period during the pandemic in which many jobs were forced to become remote, many roles may remain fully remote or hybrid long after the crisis has subsided. This despite the resistance of many organizational leaders not comfortable with managing large remote workforces and some municipalities that want workers to return to the office to reanimate central business districts. This tension requires careful study. Among the many reasons workers do not want to return the pre-pandemic norm is the burden of commuting.

It is reasonable to assume that the longer people must commute to their jobs the more likely they will find telework appealing. Commuting, depending on the circumstances, can have a beneficial effect. But commuting is not all bad. They allow for emotional, behavioral, and psychological transitions between work and private lives. It can be a period to wind down and relax. If the commuter lives in an unpleasant neighborhood, commuting to work can offer a respite. Commutes, even long ones, can afford the individual access to a better job or access to a higher quality of life. However, there are major downsides to commuting.

Long and expensive commutes tend to be unsatisfying. They can be onerous and exhausting both physically and financially. But it is important to keep in mind that not all

commutes are same even when cost and distance are held equal. The quality of the commute matters. If employers want workers back in the office or are considering where their organizations should be located, it is important to recognize the different aspects of commuting and how it may affect job-related attitudes.

In this paper, we posit that attitudes toward telework will be a function of commuting satisfaction. Commuting satisfaction is influenced by commuting enjoyment, cost, distance, strain, safety, impedance, reliability, mode flexibility, and the extent to which commuters can be productive during their commutes. Telework intentions are influenced by telework attitudes, subjective norms, and behavioral control. The next step is to evaluate and validate the model empirically in different industry domains.

REFERENCES

Ajzen, I. 1991. "The Theory of Planned Behavior," *Organizational Behavior & Human Decision Processes* (50:2), p. 179.

Allen, J., Muñoz, J. C., and Ortúzar, J. d. D. 2019. "Understanding Public Transport Satisfaction: Using Maslow's Hierarchy of (Transit) Needs," *Transport Policy* (81), pp. 75-94.

Barrero, J. M., Bloom, N., and Davis, S. J. 2023. "Why Working from Home Will Stick." Retrieved January 20, 2023, from https://wfhresearch.com/wp-content/uploads/2023/01/WFHResearch_updates_January2023.pdf

Barron, J. 2022. "The Ups and Downs of Remote Work in New York." Retrieved January 15, 2023, from https://www.nytimes.com/2022/04/12/nyregion/remote-work-new-york.html

Biron, B. 2022. "The Us Bus Driver Shortage Is 'Throwing Transit Systems into Crisis' as Big Cities Struggle to Find Public Transportation Hires." Retrieved January 19, 2023, from https://www.businessinsider.com/bus-driver-shortage-public-transit-hires-workforce-2022-7?op=1

Bosman, J., Kasakove, S., Cowan, J., and Fausset, R. 2022. "Cities Want to Return to Prepandemic Life. One Obstacle: Transit Crime." Retrieved January 19, 2023, from https://www.nytimes.com/2022/04/25/us/public-transit-crime.html

Budnitz, H., Tranos, E., and Chapman, L. 2021. "Chapter 10 - the Potential for Telecommuting to Offer Sustainable and Resilient Accessibility," in *Urban Form and Accessibility,* C. Mulley and J.D. Nelson (eds.). Elsevier, pp. 157-171.

Carvalho, M., Hagerman, A. D., and Whitacre, B. 2022. "Telework and Covid-19 Resiliency in the Southeastern United States," *Journal of Regional Analysis & Policy* (52:1), pp. 19-34.

Cerullo, M. 2023. "Like Working from Home? Some Companies Say Enough Is Enough." Retrieved January 15, 2023, from https://www.cbsnews.com/news/disney-remote-work-office-four-days-a-week/

Chatterjee, K., Chng, S., Clark, B., Davis, A., De Vos, J., Ettema, D., Handy, S., Martin, A., and Reardon, L. 2020. "Commuting and Wellbeing: A Critical Overview of the Literature with Implications for Policy and Future Research," *Transport Reviews* (40:1), pp. 5-34.

Chng, S., White, M., Abraham, C., and Skippon, S. 2016. "Commuting and Wellbeing in London: The Roles of Commute Mode and Local Public Transport Connectivity," *Preventive Medicine* (88), pp. 182-188.

Clark, B., Chatterjee, K., Martin, A., and Davis, A. 2020. "How Commuting Affects Subjective Wellbeing," *Transportation* (47:6), pp. 2777-2805.

Colbert, C. J. 2012. "Telework as Part of a Business Continuity Strategy: A Path toward Organizational Resilience." ProQuest Information & Learning, pp. 253-253.

Desilver, D. 2020. "Before the Coronavirus, Telework Was an Optional Benefit, Mostly for the Affluent Few." Retrieved January 15, 2023, from https://www.pewresearch.org/fact-tank/2020/03/20/before-the-coronavirus-telework-was-an-optional-benefit-mostly-for-the-affluent-few/

Elle, H. 2020. "The Great Rebalancing: Working from Home Fuels Rise of the 'Secondary City'," in: *The Guardian*. London, UK.

Freedman, D. H. 2022. "How Covid Opened a 'Pandora's Box' of Monkeypox, Polio and Other Diseases." Retrieved 2023, January 22, from https://www.newsweek.com/2022/09/09/how-covid-opened-pandoras-box-monkeypox-polio-other-diseases-1738153.html

Gimenez-Nadal, J. I., and Molina, J. A. 2019. "Daily Feelings of Us Workers and Commuting Time," *Journal of Transportation & Health* (12), pp. 21-33.

Goldberg, E. 2022. "Just 8% of Manhattan Office Workers Are Back Full Time, Survey Shows." Retrieved January 15, 2023, from https://www.nytimes.com/2022/05/10/business/return-to-office-manhattan.html

Henke, R. M., Benevent, R., Schulte, P., Rinehart, C., Crighton, K. A., and Corcoran, M. 2016. "The Effects of Telecommuting Intensity on Employee Health," *American Journal of Health Promotion* (30:8), pp. 604-612.

Higgins, C. D., Kanaroglou, P. S., and Sweet, M. N. 2018. "All Minutes Are Not Equal: Travel Time and the Effects of Congestion on Commute Satisfaction in Canadian Cities," *Transportation* (45:5), pp. 1249-1268.

Hoang, A. T., Nickerson, R. C., Beckman, P., and Eng, J. 2008. "Telecommuting and Corporate Culture: Implications for the Mobile Enterprise," *Information Knowledge Systems Management* (7:1/2), pp. 77-97.

Ingraham, C. 2016. "The Astonishing Human Potential Wasted on Commutes." Retrieved January 15, 2023

Jacobs, A. W., and Brailsford, J. M. 2019. "Do Work Commutes Moderate the Association between Perceived Neighborhood Disorder and Psychological Distress?," *Community, Work & Family* (22:5), pp. 589-605.

Johnson, M. 2022. "Covid, Monkeypox, Polio: Summer of Viruses Reflects Travel, Warming Trends." Retrieved January 22, 2023, from https://www.washingtonpost.com/science/2022/09/01/summer-viruses-covid-19-monkeypox-polio/

Jun, M.-J., Kwon, K.-H., and Jeong, J.-E. 2019. "An Evaluation of the Value of Time for Commuting in Seoul: A Life Satisfaction Approach," *International Journal of Sustainable Transportation* (13:10), pp. 703-709.

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., and Stone, A. A. 2004. "A Survey Method for Characteriing Daily Life Experience: The Day Reconstruction Method," *Science & Education* (306), pp. 1776-1780.

Kaplan, S., Tchetchik, A., Greenberg, D., and Sapir, I. 2022. "Transit Use Reduction Following Covid-19: The Effect of Threat Appraisal, Proactive Coping and Institutional Trust," *Transportation Research Part A: Policy & Practice* (159), pp. 338-356.

Kluger, A. N. 1998. "Commute Variability and Strain," *Journal of Organizational Behavior* (19:2), p. 147.

Koslowsky, M., and Aizer, A. 1996. "Stressor and Personal Variables in the Commuting Experience," *International Journal of Manpower* (17:3), p. 4.

Liu, J., Ettema, D., and Helbich, M. 2022. "Systematic Review of the Association between Commuting, Subjective Wellbeing and Mental Health," *Travel Behaviour and Society* (28), pp. 59–74.

Lucas, J. L., and Heady, R. B. 2002. "Flextime Commuters and Their Driver Stress, Feelings of Time Urgency, and Commute Satisfaction," *Journal of Business and Psychology* (16:4), pp. 565-572.

Lykiardopoulou, I. 2021. "Employees Hate Their Commute, Not the Office." Retrieved January 19, 2023, from https://thenextweb.com/news/employees-work-from-home-because-they-hate-commute-not-office

Lyons, G., and Chatterjee, K. 2008. "A Human Perspective on the Daily Commute: Costs, Benefits and Trade-Offs," *Transport Reviews* (28:2), pp. 181-198.

Ma, L., and Ye, R. 2019. "Does Daily Commuting Behavior Matter to Employee Productivity?," *Journal of Transport Geography* (76), pp. 130-141.

McMahon, J. 2013. "Top Eight Reasons People Give Upn on Public Transit." Retrieved January 19, 2023, from https://www.forbes.com/sites/jeffmcmahon/2013/03/06/top-eight-reasons-people-give-up-on-public-transit/?sh=3d004f5fd84c

Meymandpour, R., and Bagheri, Z. 2017. "A Study of Personality Traits, Viz., Extraversion and Introversion on Telecommuters' Burnout," *Telecom Business Review* (10:1), pp. 1-7.

Nicholls, L., Phelan, K., and Maller, C. 2018. "'A Fantasy to Get Employment around the Area': Long Commutes and Resident Health in an Outer Urban Master-Planned Estate," *Urban Policy & Research* (36:1), pp. 48-62.

Nilles, J. 1976. *The Telecommunications-Transportation Tradeoff: Options for Tomorrow*. Los Angeles, California: Krieger Publishing Company.

Nilles, J. 2018. "Jack Nilles Biography." Retrieved January 15, 2023

Novaco, R. W., Stokols, D., and Milanesi, L. 1990. "Objective and Subjective Dimensions of Travel Impedance as Determinants of Commuting Stress," *American Journal of Community Psychology* (18:2), pp. 231-257.

Page, N. C., and Nilsson, V. O. 2017. "Active Commuting: Workplace Health Promotion for Improved Employee Well-Being and Organizational Behavior," *Frontiers in Psychology* (7).

Park, G., and Evans, G. W. 2016. "Environmental Stressors, Urban Design and Planning: Implications for Human Behaviour and Health," *Journal of Urban Design* (21:4), pp. 453-470.

Raišienė, A. G., Rapuano, V., and Varkulevičiūtė, K. 2021. "Sensitive Men and Hardy Women: How Do Millennials, Xennials and Gen X Manage to Work from Home?," *Journal of Open Innovation* (7:2), pp. 106-106.

Schaeffer, M. H., Street, S. W., Singer, J. E., and Baum, A. 1988. "Effects of Control on the Stress Reactions of Commuters," *Journal of Applied Social Psychology* (18:11), pp. 944-957.

Simón, H., Casado-Díaz, J. M., and Lillo-Bañuls, A. 2020. "Exploring the Effects of Commuting on Workers' Satisfaction: Evidence for Spain," *Regional Studies* (54:4), pp. 550-562.

Sposato, R. G., Röderer, K., and Cervinka, R. 2012. "The Influence of Control and Related Variables on Commuting Stress," *Transportation Research: Part F* (15:5), pp. 581-587.

Stone, A. A., and Schneider, S. 2016. "Commuting Episodes in the United States: Their Correlates with Experiential Wellbeing from the American Time Use Survey," *Transportation Research: Part F* (42), pp. 117-124.

Stutzer, A., and Frey, B. S. 2008. "Stress That Doesn't Pay: The Commuting Paradox," *Scandinavian Journal of Management* (110:2), pp. 339-366.

Teo, T. S. H., Lim, V. K. G., and Wai Sook, H. 1999. "Assessing Attitudes Towards Teleworking among Information Technology (It) Personnel," *Singapore Management Review* (21:1), p. 33.

Thompson, R. J., Payne, S. C., Alexander, A. L., Gaskins, V. A., and Henning, J. B. 2021. "A Taxonomy of Employee Motives for Telework," *Occupational Health Science*).

Wigert, B., and Agrawal, S. 2022. "Returning to the Office: The Current, Preferred and Future State of Remote Work." Retrieved January 15, 2023, from https://www.gallup.com/workplace/397751/returning-office-current-preferred-future-state-remote-work.aspx

Williams, J. C., Andrews, O., and Boginsky, M. 2022. "Why Many Women of Color Don't Want to Return to the Office." Retrieved January 14, 2023, from https://hbr.org/2022/05/why-many-women-of-color-dont-want-to-return-to-the-office

Wingard, J. 2022. "Remote Work: Productivity up, Innovation Down." Retrieved January 15, 2023, from https://www.forbes.com/sites/jasonwingard/2022/03/17/remote-work-productivity-up-innovation-down/?sh=134aeb131b7a

Zhou, L., Wang, M., Chang, C. H., Liu, S., Zhan, Y., and Shi, J. 2017. "Commuting Stress Process and Self-Regulation at Work: Moderating Roles of Daily Task Significance, Family Interference with Work, and Commuting Means Efficacy," *Personnel Psychology* (70:4), pp. 891-922.

To Merge or Not to Merge: Analysis of External and Internal Drivers of US Mergers and Acquisitions by Sector, 1999-2018

Extended Abstract + Visualizations for Submission to the NEDSI Conference 2023

Summary: In this paper, we analyze external and internal drivers of mergers and acquisitions (M&A) in the pre-pandemic decades of 1999-2018. Our analysis uncovers patterns in M&A varying by the 11 global industry classification standard (GICS) sectors and by cycles of economic downturn and recovery. Using clustering analysis on M&A from SDC Platinum, with transactions valued over 250 million USD with publicly-traded US acquiring firms and publicly-traded or privately held US target firms, we determine similarities and differences in M&A activity by GICS sector and across economic era. The economic cycles include two major market shocks: the high-tech dot.com meltdown around 2001 and the global financial crisis in 2008-2009. We find that the distributions of M&A by GICS sectors can be grouped into three major patterns, reflecting upsurges, downturns, and resilience for recovery even following major market upheavals. Additionally, combining CEO-level data from BoardEx with firm-level SDC Platinum transactions, we have preliminary findings regarding CEO demographic background and reward drivers of M&A transactions. We focus on the period before covid-19, and implications of our research are to consider the impact of the covid-19 pandemic on current and future M&A activity, given the 2022 circumstances of global economic retrenchment and the associated decline in M&A.

Extended Abstract with Visualization and Explanations from Preliminary Findings:

The question of the correct timing for a firm merging or not merging with another firm has been intensively debated for decades, with some scholars pointing to institutional-level economic and market robustness factors as preeminent, and other scholars identifying CEO-level factors with respect to demographic background, motivations, and anticipated financial reward outcomes as strongly driving decisions to merge or not. However, these studies have not adequately addressed that each cycle of economic downturn and recovery has both commonalities and differences that must considered as external drivers of market forces *alongside* internal (to the firm) demographic, motivational, and reward drivers at the CEO level. The separation of consideration of these external and internal factors has created a conundrum. Isolating the economic and market circumstances from the dynamics at the leadership level inside the firm *and* not looking at each economic downturn and recovery as both unique and convergent with previous such incidents has led to a gap in the completeness of our understanding around the pursuit of mergers and acquisitions (M&A) and the decision making for merging or not (here using the term merger as synonymous with the term M&A). Additionally, the actual occurrence of the transactions must be taken into consideration.

Our paper addresses patterns and dynamics of M&A occurrence with attention to three instances of economic disruption—the dot.com bubble, the global financial crisis, and,

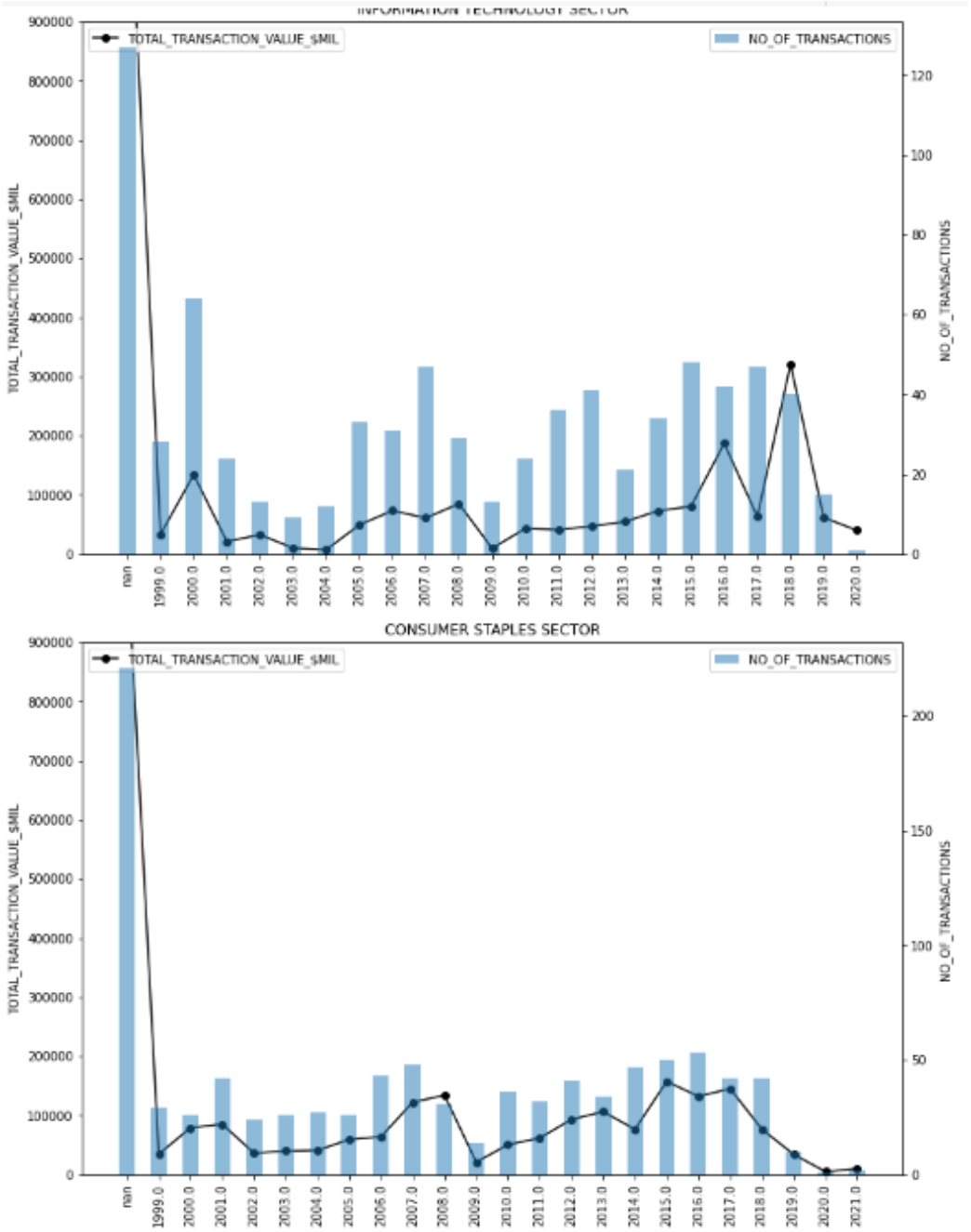
for future evaluation, the covid-19 global pandemic—while also taking potential CEO level background and outcome factors into consideration. Specifically, drawing on data from the SDC Platinum and BoardEx databases, we examine US M&A valued over 250 million USD in the 11 global industry classification standard (GICS) sectors (Figure 1) from 1999-2018 for n=14,405 M&A and approximately n=500 CEOs. Using clustering analysis, we can demonstrate groupings into patterns of transactions before and around the time of the dot.com bubble (1999-2001) and before and after the global financial crisis (circa 2008-2009) to show that M&A decreased and then resumed across sectors, although to different degrees for each sector and for each crisis and recovery Figure 2). In addition, our initial analysis of our CEO-level data (Figure 3), although more limited (Figure 4), indicates merging to be a predominantly male CEO-driven activity with inferred egoistic, reputational, and power rewards as well as longer-term wealth-building effects, as direct non-contingent compensation does not markedly change for CEOs in the aggregate before and after M&A (Figure 5).

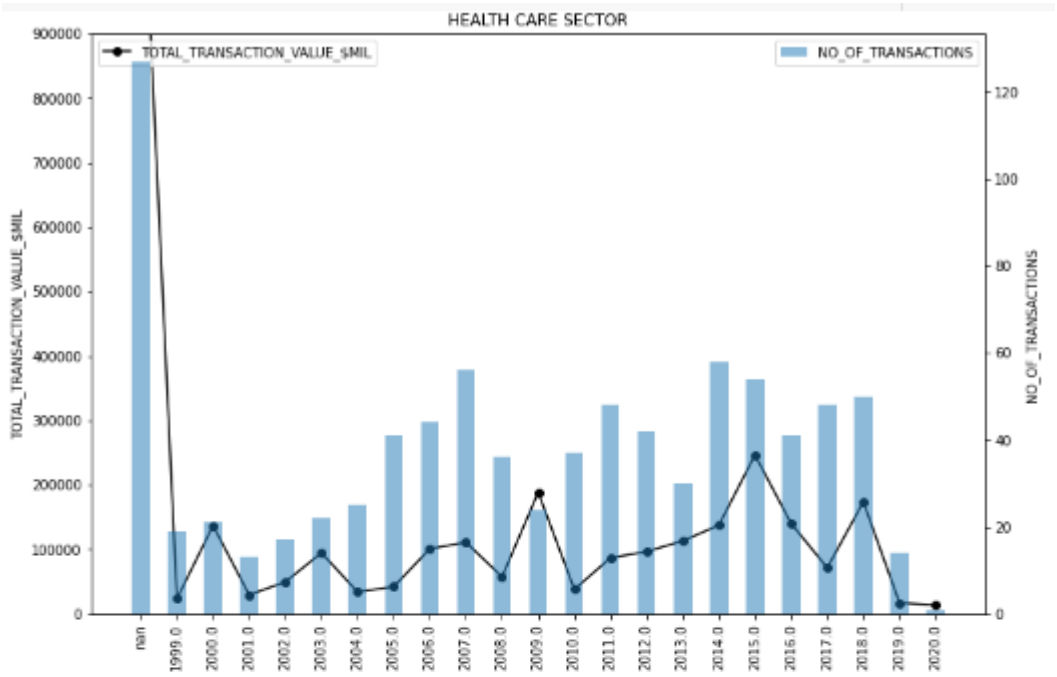
Our findings have implications for understanding the covid-19 global pandemic associated economic and market downturns (circa 2020-present). In essence, we argue that both economic and market forces *and* CEO level factors should be taken into consideration in better comprehending merging or not merging. As our present universe of data encompasses only US publicly traded acquirers and US targets, with M&A valued over 250 million USD, representing the 11 GICS sectors, and with a more limited amount of CEO level data available, our findings also have implications for collecting additional US CEO level data and pursuing further studies involving M&A internationally. In conclusion, by closely examining the convergence of economic and market factors with CEO demographic background, motivational and reward factors, our research emphasizes the importance of studying firms in conjunction with their leaders and sheds light on the neglected issue of the previously excessive partitioning of M&A contributing factors into external and internal domains. We have provided some Tables and Figures in preliminary illustration of our research setting.

Figure 1: Mean M&A valuation (in millions of USD) by GICS sector during 1999-2018 (deal size ≥250m USD): The financials, consumer discretionary, and industrials sectors experienced the highest average M&A activity from 1999-2018. We also provide the IT sector graph by way of example of a sector with moderately high average M&A activity.

gics_target_ind_sec	gics_target_ind_sec
Communication Services	973
Consumer Discretionary	1680
Consumer Staples	1375
Energy	1155
Financials	2386
Health Care	1058
Industrials	1500
Information Technology	1026
Materials	1010
Real Estate	909
Utilities	1333

Figure 2: M&A events in GICS sectors, 1999-2018: All sectors experienced notable M&A activity from 1999-2018, with intermittent declines and rebounds corresponding to the two major economic shocks of the time: the high-tech, dot.com meltdown circa 2001 and the global financial crisis circa 2008-2009. Reduced rates of M&A activity occur in the approximately 1-3 years post economic and market shock. The patterns in each sector may appear roughly similar in upswings, downturns, and recovery around each shock, but clustering analysis can reveal a more nuanced distinction and grouping in the patterns of M&A activity and resilience by sector and shock.


TOTAL_TRANSACTION_VALUE_$MIL
NO_OF_TRANSACTIONS
CONSUMER STAPLES SECTOR


HEALTH CARE SECTOR
TOTAL_TRANSACTION_VALUE_$MIL
NO_OF_TRANSACTIONS
TOTAL_TRANSACTION_VALUE_$MIL
NO_OF_TRANSACTIONS
900000
800000
700000
600000
500000
400000
300000
200000
100000
0
120
100
80
60
40
20
0
nan
1999.0
2000.0
2001.0
2002.0
2003.0
2004.0
2005.0
2006.0
2007.0
2008.0
2009.0
2010.0
2011.0
2012.0
2013.0
2014.0
2015.0
2016.0
2017.0
2018.0
2019.0
2020.0

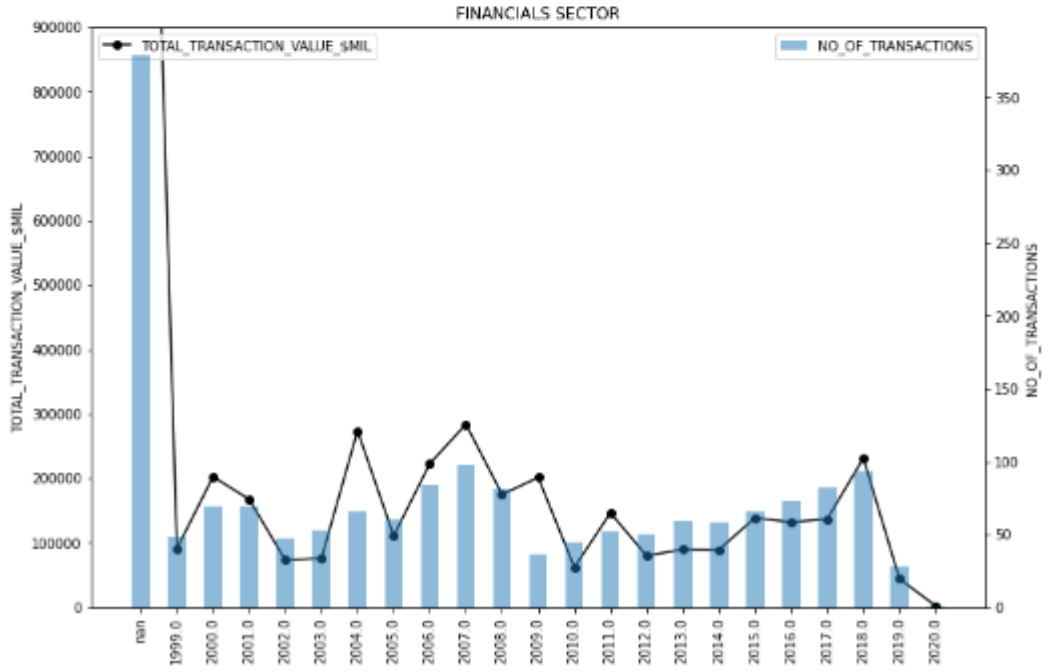

FINANCIALS SECTOR
TOTAL_TRANSACTION_VALUE_$MIL
NO_OF_TRANSACTIONS
TOTAL_TRANSACTION_VALUE_$MIL
NO_OF_TRANSACTIONS
900000
800000
700000
600000
500000
400000
300000
200000
100000
0
350
300
250
200
150
100
50
0
nan
1999.0
2000.0
2001.0
2002.0
2003.0
2004.0
2005.0
2006.0
2007.0
2008.0
2009.0
2010.0
2011.0
2012.0
2013.0
2014.0
2015.0
2016.0
2017.0
2018.0
2019.0
2020.0

Figure 3: Distribution of CEO age by GICS sector, 1999-2018: In general, the CEOs in the sample are in their mid-to-late 50s to early 60s.

gics_sector	Age
Communication Services	61.0
Consumer Discretionary	57.0
Consumer Staples	57.0
Energy	60.0
Financials	60.0
Health Care	53.0
Industrials	57.0
Information Technology	51.0
Materials	56.0
Real Estate	54.0

Figure 4: Frequency of M&A by GICS sector during 1999-2018 with CEO data available: Data were not available for all CEOs for firms in all sectors. Due to missing data, our present findings are preliminary. The dataset currently has the highest representation of CEOs in the financials, industrials, and energy sectors. The first two of these sectors are also two of the sectors of the highest M&A activity, suggesting the validity of our subsequent observations of male CEO predominance in M&A activity and a discernible lack of compensation differentiation for CEOs pursuing or not pursuing M&A across the two decades of our time period.

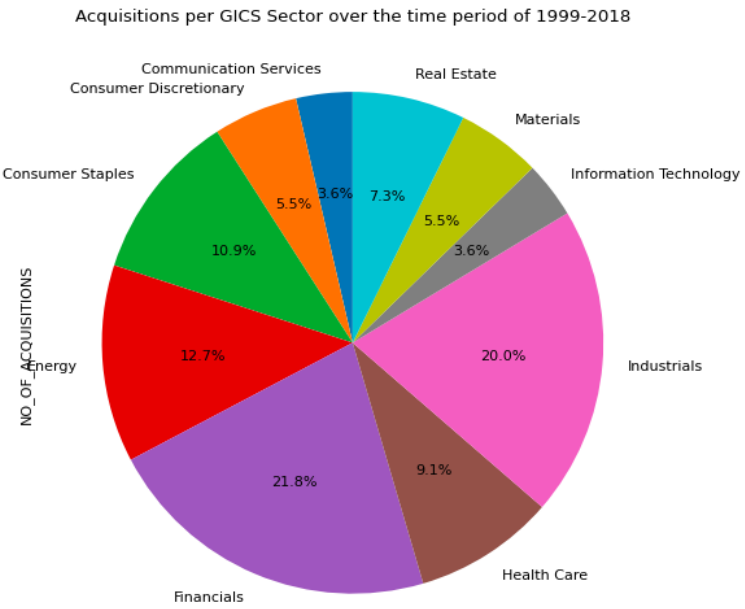


Figure 5: Mean salaries of female and male CEOs pursuing and not pursuing M&A: CEOs in firms that have pursued M&A do not consistently have a higher or lower salary level than CEOs in firms not experiencing M&A, but there is an absence of female CEOs in firms experiencing M&A.

`ceo_with_ma_sal_freq`

	gics_sector	Gender	Salary
0	Communication Services	M	1965.30
1	Consumer Discretionary	M	850.00
2	Consumer Staples	F	NaN
3	Consumer Staples	M	1121.47
4	Energy	M	783.89
5	Energy	n.a	NaN
6	Financials	M	964.23
7	Financials	n.a	NaN
8	Health Care	M	1621.28
9	Health Care	n.a	NaN
10	Industrials	M	957.15
11	Industrials	n.a	NaN
12	Information Technology	M	NaN
13	Information Technology	n.a	NaN
14	Materials	M	1168.88
15	Materials	n.a	NaN
16	Real Estate	M	NaN
17	Utilities	M	0.00

`ceo_without_ma_sal_freq`

0	Communication Services	F	NaN
1	Communication Services	M	846.53
2	Communication Services	n.a	NaN
3	Consumer Discretionary	F	2075.00
4	Consumer Discretionary	M	817.22
5	Consumer Discretionary	n.a	NaN
6	Consumer Staples	F	894.73
7	Consumer Staples	M	1163.63
8	Consumer Staples	n.a	NaN
9	Energy	F	NaN
10	Energy	M	959.88
11	Energy	n.a	NaN
12	Financials	F	925.00
13	Financials	M	836.63
14	Financials	n.a	NaN
15	Health Care	F	NaN
16	Health Care	M	952.74
17	Health Care	n.a	NaN
18	Industrials	F	1562.80
19	Industrials	M	944.15
20	Industrials	n.a	NaN
21	Information Technology	F	1214.00
22	Information Technology	M	709.96
23	Information Technology	n.a	NaN
24	Materials	F	NaN
25	Materials	M	957.47
26	Materials	n.a	NaN
27	Real Estate	F	NaN
28	Real Estate	M	568.36
29	Real Estate	n.a	NaN
30	Utilities	F	NaN
31	Utilities	M	986.35

Trusting the Fall: Repetitive Obstacle Performance Evaluation System (ROPES) Course as Organization Development and Change Intervention

2023 Northeast Decision Sciences Institute 52nd Annual Conference
https://nedsi.decisionsciences.org/conference/

Abstract

The need to increase organizational effectiveness is an ever-present challenge faced by leaders. Building interpersonal and intrapersonal skills of teams within the organization is a way to increase effectiveness to meet dynamic opportunities in pursuit of strategic vision and alignment. This paper addresses the use of ROPES courses by organizational development practitioners as an intervention strategy to achieve greater synergy. The central question relates to the effectiveness of such an intervention and if it is a meaningful and effective use of time and resources. A literature review examines the historical use of ROPES courses as a means of developing teams, the advantages and disadvantages of use, and opportunities for successful implementation. The key takeaway is that with a methodical approach that conforms to the needs of the organization, a ROPES course-based team development intervention can be successful but only if the lessons learned on the course are properly reflected on and implemented in the organization.

Keywords: team building, corporate retreats, group dynamics, effect of team building in business, ROPES course facilitation, “benefit of shared experience”

Introduction

Are people ready to fall? Do they trust the process? Will they let themselves go towards perceived risk with the goal of achieving something greater? These basic questions relate to the easily understood and known activity often encountered in a team building event – the trust fall. Referenced as early as 1984 in Karl Rhonke's *Silver Bullets: A guide to Initiative Problems, Adventure Games, Stunts and Trust Activities* this simple activity demonstrates the interconnected reliance on a group by an individual to face an unknown, at personal risk, amid a defined vision and goal. There is a parallel here to an organization, one which with proper utilization of team development activities can enhance the capabilities of an individual and a group. These activities are not just games, but when properly led, facilitated, and reflected upon form a basis for growth while providing a cohesive element to group development within an organization.

A trust fall activity is one of a plethora of events that a group can engage in. These activities can take place in an office setting, or on a well-constructed and defined ROPES course. "ROPES" is an often-used acronym for a variety of activities that involve physical and mental challenges which as best accomplished by leveraging the collective strength of a team. Decision-making and participation are the hallmarks of the experience as team members engage in the "Repetitive Obstacle Performance Evaluation System" (Daniels, 1997, p. 239).

Activity can be scaled from very vigorous to sedentary yet still achieve a great deal of intellectual engagement as well as group participation. Though generally perceived as fun, these activities have a greater meaning, or should. In particular, such experiences may be beneficial for teams with high task interdependence or workgroups seeking to improve cohesiveness, trust, or a

sense of efficacy" (Eatough et al., 2015, pp.82-83). The "trick" is to reflect, to process the experience and then apply the lessons learned on the course into the office.

Building teams, stepping out of comfort zones and growing as a group are all fundamental aspects of an organization that seeks to thrive in a competitive environment. Cost considerations coupled with market competition weigh upon the minds of managers and leaders as they seek to tackle the challenges of their profession. Team building, or group development, is integral to achieving success, as a unified team will achieve greater than a disparate collection of individuals doing their own thing. This concept is in line with the first-generation trend of Organization Development whereby there is a focus on change at the individual level. Despite a recent shift in practitioner focus to a broader focus of an organization's attention to and alignment with its external environment, also known as the second generation, there still is considerable emphasis placed upon the employee through pervasive self-managed work teams who are given autonomy and ownership of the work they accomplish (Anderson, 2020, p. 37).

The ROPES course is an opportunity to challenge people in a group setting with the goal of translating lessons learned into transferable skills. The development of the team is best defined as, "the activity of attempting to improve a work group's effectiveness at doing its work, maintaining the relationships of its members and the team's contributions to the wider organizational system" (Coughlan, 1994, p.21). However, without proper facilitation of the experience and follow-on support within the organization the lessons learned will likely fade away. This negative experience can affect not only the group but also the perception of the ROPES course as an intervention to build teams. The group may hold the failures as proof that the task at hand is too lofty, or that they are not as cohesive as needed to be. Likewise, the use of the experience is cheapened through improper facilitation thus negating the potential positive

impacts for other groups. Thus, it becomes incumbent upon all, the group as well as the facilitator, to define the goals of the intervention, to give an honest and complete effort, and to carry on the lessons learned throughout the lifecycle of the group's future.

Statement of the Problem

Organizations need tested teams comprised of effective leaders and followers who can think abstractly to solve unique and emergent issues. Effective groups are those that have trust built upon communication, shared vision, and the willingness to be vulnerable with one another. To be competitive, organizations will continue to try new interventions to gain an advantage. Having fun is a quality indicator of employee satisfaction but not always a mechanism to improve an organization's efficiency. The use of ROPES courses can be the gap between fun and productivity, but only if there will be a return on the investment, a return brought forth by proper intervention coupled with long-term support.

Justification for Study

The justification for this study is two-fold. Firstly, the need to address the issue we have articulated above as the statement of the problem and the Interest to the authors borne out of lived experiences and observations. The justification of our study in our attempt to address the statement of the problem include:

1. *Improving teamwork and communication:* ROPES activities can be effective in promoting teamwork and improving communication within a team, which can lead to improved performance and effectiveness.

2. *Building trust and cohesion:* ROPES activities can help to build trust and cohesion within a team, which can foster a more positive and collaborative work environment.

3. *Enhancing leadership skills*: ROPES activities can help to develop leadership skills and encourage individuals to take on leadership roles within the team.
4. *Promoting team development*: ROPES activities can be used as a team development intervention to address issues such as conflict, communication breakdowns, and lack of trust within the team.
5. *Improving employee engagement and motivation*: ROPES activities can be a fun and engaging way to motivate and engage employees, which can lead to improved job satisfaction and retention.

Overall, a study on ROPES for organization development could provide valuable insights into the effectiveness of this type of intervention in improving teamwork, communication, trust, and performance within an organization.

This topic is of interest to the authors for a variety of reasons and demonstrates a lifelong commitment to the development of self and others. The lead author of this article has expertise in Outdoor Recreation Management. The lead author was a ROPES course facilitator. Working with well over one hundred various groups, he gained an appreciation for as well as an understanding of the benefits as well as shortcomings of using a ROPES course experience to develop groups. The observations were that some groups merely participated in the activity because it was fun. While others utilized the experience to bring themselves closer, working through issues present within the group, and to thrive. The observation was that most of what the lead author of this manuscript was doing as a facilitator at the time was very transactional as the groups utilized his expertise on the course but then left without additional input of how their shared experiences led to further development. As the story goes, it wasn't until he worked

alongside another facilitator who was consulting with a local grocery store chain management team. Over the course of several weeks this team would conduct off site retreats at the Lead Author's facility to engage in activities. Then they would process the experiences in terms of companywide objectives and projects that they were working on. Through this experience the lead author saw the long-term value of the activities, of the practical application of experience into action.

Research Questions and Method of Inquiry

We shaped our inquiry and review around the following questions: What is the conceptual basis for using ROPES activities as a team-building intervention for organization development?

1. How do ROPES activities align with theories and models of team development and performance?
2. What are the key mechanisms through which ROPES activities may influence teamwork and performance within an organization?
3. How do contextual factors, such as the culture of the organization or the specific needs and goals of the team, influence the effectiveness of ROPES activities for organization development?
4. How can ROPES activities be effectively designed and implemented for organization development?

In addressing these questions, we utilized a variety of available online resources. Initial searches were focused on the psychological impact of team building, specifically focused on the formation and bonding of groups. Next, searches focused on the business impact of group development to see what research has been done on the topic and to add support to the problem

statement. Google Scholar, ProQuest, EBSCOHost, ERIC, and JSTOR were databases used. We followed a multidisciplinary integrated literature review approach in our attempt to transcend the mere summation of the literature but contribution to the extant literature in the fields of leadership and organization development (Marczak & Yawson 2021; Osafo & Yawson, 2021). The literature search covered a variety of academic fields, including psychology, organization development, human resource development, organizational behavior, and other related social science fields.

Literature Review

Getting Roped In

The focus of our research is on the merits of the ROPES course as an intervention tool to be utilized when developing teams. The multifaceted and interdisciplinary nature of the collection of events engaged by teams over the course of a training event serves to further the group's development. This development is further enhanced through the use of reflection and then the application of lessons learned in a work environment. As the literature suggests, "a ROPES course requires dependence on others for success, yet is influenced by participants' own knowledge, skills, and abilities. Such a challenge is expected to lead to inter- and intrapersonal learning and growth" (Eatough et al., 2015, p. 65). The ROPES course is therefore a laboratory setting where the effects of the experience can be used by both the organizational consultant as well as the group to foster desired outcomes.

Eatough et al. (2015) present a review of available literature to identify instances of acute stressful events that a work group may engage in. This, done in a controlled environment, illicit the psychological effect of trust, cohesion, group and self-efficacy. "Across the literature, average effects of ROPES course interventions on psychological outcomes (regardless of sample

type or variable of interest) have been reported at d = .45. Some outcomes such as interpersonal dynamics have effect sizes greater than .60" (Eatough et al., 2015, p. 68).

Eatough et al. (2015) found that while groups benefit from the ROPES course intervention, it is actually the individual who has not previously been exposed to stressing events where they needed to rely on group support that gained the most value from the experience. As such,

> "workgroups should consider employing training techniques involving ROPES course challenges or similar stress-inducing activities where group reliance is necessary to foster improvement in group- and individual-level outcomes. In particular, such experiences may be beneficial for teams with high task interdependence or workgroups seeking to improve cohesiveness, trust, or a sense of efficacy" (Eatough et al., 2015, pp.82-83)

It is especially true for groups with a high task interdependence or in groups that seek to develop further individual or interpersonal skills which can benefit the group as a whole. However, there is a methodical approach to the intervention that should be followed and the intervention should not be seen as merely a fun event nor one that is done in isolation, without a defined objective or consistent follow-up post-intervention. Addressing the events, the outcomes, and the correlation to work enshrine the experience as a developmental tool.

Elevating Leadership Development Practices to Meet Emerging Needs

We attempt to address the unique challenges that face leaders, specifically focused on the leaders' capability to lead in volatile, uncertain, complex and ambiguous (VUCA) environments as demonstrated through vertical and horizontal development (Chesley et al., 2020). Traditional leadership development occurs along the horizontal axis but ROPES course interventions offer unique opportunities for vertical development. The horizontal and vertical development are depicted in Table 1.

Table 1: Traditional Leadership Development and ROPES Course Intervention

Horizontal Development	Vertical Development
Traditional leadership development programs	An emerging, evidence-based set of practices
Focus on building skills and adding competencies	Focus on Growing the mindset agility that leads to strategic thinking
Do what we already know how to do better	Learn how to lead in a dynamic, complex and uncertain world

Adapted from Chesley et al. (2020)

A research conducted by directly interviewing nineteen internal leaders responsible for leadership development from fifteen large organizations across a wide spectrum of industries demonstrated that data gathering supported a vertical approach as it tended to not only understand the needs of the organization but also the unique characteristics of interpersonal capability that leaders inherently possess (Chesley et al., 2020). Specifically, the vertical approach emphasized, "following up with individuals, using the data as a way to start conversations that provide further feedback on both current behaviors and planned growth areas" (Chesley et al., 2020, p. 183).

Chesley et al. (2020) study, enshrine the Organizational Development Consulting Model whereby, "emphasis is placed not on the receiving of expert knowledge but the integrating of expert knowledge" (Chesley et al., 2020, p. 185). Furthermore, the experiential nature of a ROPES course intervention supports the vertical approach by introducing experiences that emphasize personal growth in a dynamic, often high-stakes, environment.

Exploring Social Constructions of Followership

Creswell (2018) utilize a grounded theory and a cross-sample interviewing approach to gain an understanding of perceptions of followership from a variety of respondents in multiple fields in the United States and Canada. The intent was to build upon the limited body of data supporting followership in so much as it focuses on the nuanced perspective of what

followership means to the individual. In doing so Creswell (2018) conducted semi-structured interviews and cross-sectional qualitative sampling to obtain a broad range of experiences. The methods are conducive to answer the studies stated question. The methods achieve the desired goal of ascertaining a broad perspective irrespective of a particular occupational field, a particular seniority within a profession, or a geographic boundary. Further, the focus is not on how followership fits into the broader leadership domain but rather on the perception of followership. Thus, the individual perspective is important as it underscores bias, perception, or feeling associated with followership. In doing so there is clarity to the perspective which can be utilized to further additional research or to shape the delivery of followership as a topic.

Though the Creswell (2018) study has merit, there are areas of weakness. It seems like this was a well-done endeavor as the study had an assumption that was tested and utilized a variety of control measures to ensure the validity of the data. An area of weakness is related to the size of the studied population and the varied fields. The intent was to remain generally broad, yet while the research achieves this goal it yields little value to anyone needing specific information and could be subject to a variance if the sample size were larger. Meaning, if the sample were larger, and respondents answered in different ways, then the results could alter accordingly.

Another, and similarly related weakness, is the seemingly haphazard selection of occupational specialties of the studied population. As the authors state, “participants represent a variety of public and private sector organizations and industries including education, high-technology, financial services, mass media, consulting, and health care. The average tenure among participants with their current organizations was approximately 9 years (s.d.=6.1) and average tenure working with their current supervisor was 3.5 years (s.d.=2.69)” (Carsten et al.,

2010, p. 547). We take aim at this breadth of occupational focus, coupled with the minute sample size, to question the overall validity of the research. Yes, in research sample sizes are small. Yet, we are not sure the desired outcomes are truly achieved in such a small and sporadic size, as done here. We contend that fixating on a particular field or increasing the sample size would have been value-added as it promoted a broader perspective. The broader perspective or occupational specialization with a larger sample size may also achieve the same effect.

Outdoor Adventure and Organizational Development

Daniels (1994) conducted a study on the outdoor adventure and organizational development using ROPES course as an OD intervention. The research is dated but provides a baseline understanding of available studies performed in the later part of the past century. It advances the body of available knowledge by applying the concepts to a controlled group partaking in leadership development. Data collection, in the form of surveys, was utilized to establish a baseline understanding of the control group. Teams were then established following a T-Group model. The teams were then led through a workshop that utilized concepts of Tuckman's model of group development. The workshop included the use of a ROPES course where, "activities were chosen with regard to the importance of work problems identified by the participants in the pre-intervention survey" (Daniels, 1994, p. 243).

Methodical planning prior to, application of data collected during the intervention, and follow-up through the implementation of new teams in the organization after the event yielded significantly measurable behavioral modifications. Though, "while this does not prove that the workshop was solely responsible for the increase in cohesiveness, it does provide supporting evidence for the workshop's effectiveness" (Daniels, 1994, p. 247). Using a ROPES course was described as one of the most profound experiences in the lives of some of the participants and

lends credibility to not only the ROPES course as a team-building intervention but also to the careful application of the event as a holistic tool to be utilized by organizational development practitioners.

Meaningful Involvement Opportunities in ROPES Course Programs

Haras et al. (2006), conducted a study on the meaningful involvement opportunities in ROPES course programs. The study presents a novel approach to the facilitation of the ROPES course experience as a mechanism for team development. Traditional use, developed by organizations like Outward Bound and Project Adventure, introduced the benchmark standard for participant involvement focused on Challenge by Choice. Modifying this concept is the model of Inviting Optimum Participation. Though both models focus on introducing ROPES course events that challenge the individual, and group, in an effective way, ultimately the individual chooses the level of participation they will engage with. Therefore, the group experience may differ based on the individual participation. Table 2 outlines the differences between the two involvement opportunities.

Table 2: Differences between Challenge by Choice and Inviting Optimum Participation

Challenge by Choice	Inviting Optimum Participation
Sequential tasks congruent with participant goals, needs, and abilities	Deliberately designed challenge environment
Enhance self-efficacy and build group cohesion	Balance of participant individual characteristics
Active engagement in ROPES Course Program	Active engagement in ROPES Course Program
Novel and stimulating physical environment with support of peer group	Single inclusive activity design with multiple options and levels of challenges
Group support encourages individual participation	Positive interpersonal and intrapersonal growth and development
The facilitator leads prescribed flow of events	The facilitator provides a menu of opportunities, and the group chooses their own experience

Adapted from Haras et al. (2006)

As depicted in Table 2, the main difference between the two intervention modalities lies within the sequencing of challenging events. Though both offer many similarities, the Inviting Optimum Participation method is congruent with the Organizational Development Consulting Model as the facilitator is a part of the decision-making process but presents options which the group can use to determine their experience in a way which encourages maximum participation. "In other words, programs are successful when participants experience involvement they find meaningful. In ROPES course literature, meaningful involvement occurs when participants play a role in the activity that contributes to achieving the group's task" (Haras et al., 2006, p. 343). Furthermore, the combined effects support Csikszentmihalyi's theory of Flow whereby "Flow: focused attention on an achievable and enjoyable activity; intentional behavior chosen to give purpose to action; and harmonious inner thoughts and feelings" (Haras et al., 2006, p. 343).

Though both methods are effective in the implementation of the ROPES course as a team-building intervention, the Inviting Optimum Participation model supports further participation, involvement, and empowerment of teams as a method of supporting social construction. The organization development practitioner is able to standardize the effects of the intervention across multiple groups in a more thorough manner thus supporting continued opportunities for reflection as well as implementation after the intervention.

Discussion

We have focused on organizational development from the socially constructed point of view. If people are the heart of an organization, and that their continued support is needed to achieve effective organizational change, then key constituencies within the group would need to be aligned for change to positively occur (Andreson, 2020, pp. 90-91). Thus, there is a great reliance on the importance of groups and teams (Andreson, 2020, p. 52) to be strengthened, aligned, and guided throughout the process. Further analysis can be applied to the organizational

development practitioners utilizing the ROPES course concept as a way to develop effectively functioning teams. The focus of this would be to highlight the long-term impacts of teams who are able to enhance their capabilities through a ROPES course experience, the associated reflection, and recall of the experience towards achieving meaningful development.

The use of the ROPES course as an intervention towards team building is not novel but the implementation of effective facilitation is critical to the achievement of desired outcomes. The value of adventure-based experiential training opportunities cannot be understated as a mechanism for leaders to develop along a vertical model and for followers to have a significant impact on the development of teams. Prior to the use of the intervention is the need for the organizational development practitioner to work with the organization to define what changes, attributes, or goals are sought. With a clear purpose, the intervention can be implemented in a meaningful way. This cannot be a standalone event, but rather a part of a holistic approach, an approach that focuses as much on the experience as well as on the continued post-experience implementation of the dynamic lessons learned into the strategic vision.

In addition to the process, the participating members of the ROPES course intervention should have buy in, or acceptance of, the intervention. Maximum participation on the ROPES course is as essential as maximum participation within the team in a business setting. Though physical and psychological limitations may present challenges, the organizational development practitioner can mitigate a diminished response through the use of the Inviting Optimum Participation model. This, coupled with effective mentoring and reflection will ensure an organization’s vision is transferred to the individual and that the individual is able to utilize their development on the ROPES course to further support the organization.

Theory to Practice Application

The deep woods of Maine intersect the evolution of the ROPES course as a part of the Outward Bound program with the historical National Training Laboratory (NTL) and the advent of modern organization development (Wagstaff) (Anderson, 2020). Though each discipline has grown, organization development has seen the greatest maturation throughout the ages. The NTL T-group can be considered the baseline for ROPES course-based team development as the early programs were designed for small groups, facilitated by trained personnel, who would "create interpersonal change by allowing individuals to learn about their own and other behaviors so that this education could translate into more effective behaviors when group members went home" (Anderson, 2020, p. 26). Though T-groups have largely been replaced by contemporary theories or models, the evolution of the concept remains prevalent throughout time. That is, a shared experience in a leadership laboratory can be used in an organizational setting.

Team building interventions

Team building interventions can generally be divided into two separate categories: team development or team interventions (Anderson, 2020). Though the names are often used interchangeably within the ROPES course community, the difference in the type of group will dictate which primary intervention is used.

Team development interventions are activities or processes that aim to improve the functioning and effectiveness of a team (Salas et al 2018). These interventions can be used to address issues such as conflict, communication breakdowns, and lack of trust within the team. Examples of team development interventions include team-building activities, team training, and team coaching.

Team interventions, on the other hand, are activities or processes that focus on improving the overall performance of the team. These interventions may involve changes to the team's

structure, processes, or goals, and are often aimed at increasing productivity, improving efficiency, or enhancing innovation. Examples of team interventions include reorganizing the team, introducing new technology or tools, and setting new performance targets.

There is a wealth of literature on team development and team interventions, including research on the effectiveness of different approaches and the factors that influence their success (Lacerenza et al 2018). For example, a study by Griffin and Moor (2004) found that team-building activities can be effective in improving communication, trust, and teamwork within a team, while a review by Pearce and Sims (2002) found that team training can be effective in improving team performance. However, the effectiveness of these interventions can vary depending on the specific needs and goals of the team, as well as the context in which they are implemented.

The purpose of the challenging event is to illicit a response within the team, thereby reflection can be utilized to produce a tangible result, often awareness, of the interpersonal dynamics at play. Thus, tailoring the experience to the type of group is essential to an effective delivery of the intervention as opposed to a “cookie cutter” style of programing without coherent structure or focus.

Looking more closely at the two types of interventions highlights the difference observed with a newly formed group, team development, as opposed to a group that has been working together for some time and may be in the midst of working through barriers to increased performance capability, team interventions (Anderson, 2020). Though generally expected to return positive results, several researchers suggest that team-building interventions fail to achieve any lasting effect. In fact, “team building has a negative connotation for some clients because it cannot be denied that team-building activities often fail to achieve their objectives (Anderson,

2020, p. 261). There is also older research that suggests a pattern of regression for groups after the completion of an intervention. However, if the purpose of organizational development is to change a behavior, an attitude, and generally improve overall performance then the development of communication, trust, problem-solving, and cooperation skills are essential (Daniels, 1997).

Developing the team building intervention is a careful process that should be conducted in concert with the requesting organization by the organizational development practitioner (Mandal, 2022). Among other elements, there should be an awareness of the group's leadership and followership styles, as well as the interrelated connection to task completion. Utilizing Kelly's 1992 model of followership with Hersey and Blanchard's Situational Leadership Theory can be effective. This awareness allows for a more thoughtful approach to not only the lifecycle of the training but also to the roles of leaders and followers within the group since, "followers and leaders work together better when they are comfortable with each other, and value congruence is one way to achieve common ground" (Bjugstad et al., 2006, p.307).

Participation, Involvement, and Empowerment

Within the context of team development, it is essential for members of the groups participating in the interventions to be engaged in the process. Far more important than merely being present, the participation must be authentic for the best results to be achieved. To achieve this, organizational leaders "must develop an environment in which authentic participation is possible and organizational members can have a legitimate impact (Anderson, 2020, p. 53). Then the group members need to have the time, space, and guidance from the organizational development practitioner, to feel comfortable in the process. The participation transfers to feelings of community whereby group members will go to greater lengths to perform for one another. For instance, "According to the effort justification component of cognitive dissonance theory, it is plausible that extra effort expended to cope with stressors presented to one's

workgroup may result in increased liking of the workgroup to reduce cognitive dissonance" (Eatough et al., 2015, p. 70).

Support to Social Construction

Leaders exist because there are people within their sphere of influence, within the defined group, who follow. The study of followers is developing, that is, the majority of research on leadership is focused on the principal leader figure. However, the followers are not only participatory in the process, but they also serve a fundamental role in the entirety of the exchange. Within follower research, the focus has primarily been on the interaction between leader and follower (Carsten et al., 2010, p. 543). Yet, there is an underlying dynamic among the followers which needs further exploration, an uncharted area ready for further analysis. Multiple viewpoints are required yet the current orientation is toward the leader, so much so that the research focus is on the relationship of the leader and the follower.

Understanding the dynamics of transitional points of view associated with the leader and follower paradigm is important to the implementation of ROPES course-based team development interventions. Within the social construction perspective is an emphasis on the active role of all members in the creation of the organization (Anderson 2020). Communication also becomes fundamental to the intervention, to future change, and generally to the success of the group.

The focus of social construction within the ROPES course intervention to team development is on the participatory nature of the group members on the outcome. Anderson notes that "change can best be accomplished when organizational members have the opportunity to work together to define new practices" (Anderson, 2020, p. 92). Within the ROPES course experience, the focus is on the process rather than the outcome.

The role of the consultant using the Organizational Development Consulting Model

The experience that comes from the ROPES course intervention can enrich the human systems of an organization. Thus, the role of the consultant is best defined through the lens of the Organization Development Consulting Model as consultants are hired as experts in process consulting rather than in specific subject area expertise (Anderson, 2020). The consultant acts as a facilitator, as a guide who can aid the client in discovering the solution internally but through the process of self-discovery and group development. As mentioned in prior application topics, "facilitators may do well to encourage participants to focus less on "solving" the challenge and more on applying other efforts that will assist with the challenge, such as communication and encouragement of others" (Eatough et al., 2015, p. 70).

One of the greatest missed opportunities discovered, when it comes to the application of ROPES course interventions, is the lack of consistent engagement throughout the planning, execution, and post intervention phases. The Organizational Development Consulting Model reduces the chance of missed opportunities through client engagement which begins early and is often relied upon throughout. "The client's active participation is necessary to ensure that once the consultant leaves the engagement, the intervention can be a sustained change. Without a client's active participation, the results are likely to be short-lived" (Anderson, 2020, p. 108). Again, the research demonstrates the interconnected relationship of organizational development fundamentals using the ROPES course as a vehicle to achieve greater success but only if certain basic elements are consistent.

Observation as a form of data gathering

Throughout the intervention process, an Organization Development practitioner may utilize several forms of data-gathering tools. However, the use of direct observation is the most beneficial and accurate measure to utilize while engaging in ROPES course activities. Despite

limitations of observation, specifically the Hawthorne Effect, among participants, the application of a ROPES course intervention can occur over several hours, or even days. Thus, the practitioner can observe interactions among the group over time and allow the practitioner to assess data in real-time (Anderson, 2020).

The follow-through requirement of using the intervention to achieve meaningful outcomes invites the debate about the type of consultant to be used. In some respects, it is best if a trained ROPES course facilitator is present to facilitate the experience but does so in concert with an in-house consultant. Thus, the observations made during the intervention can be studied over time, and in different situations. The in-house consultant would also be preferred in many cases as they will likely have a greater rapport with the client which can aid in processing the intervention and developing future solutions based on lessons learned.

Implication for Future Research

Research within the field of organizational development has increased significantly and changed throughout time since the inception at the NTL. Specific research addressing the use of ROPES course, or other adventure-based intervention modalities, continues to be limited as compared to other aspects of organizational development and change. The available body of research has relied on classic themes but falls short of truly understanding the dynamic effects of the current world.

Within the next few years, the mantle of leadership will be passed through generations as the Baby Boomer generation transitions to their next phase of life. Thus, the research focused on teams will need to evolve to incorporate the new era of leaders and followers, complete with their uniquely different views on the relationship with work as well as with their peers. Value in

the experience, as well as in the work performed, will need to be communicated. Fresh concepts in team development will likely evolve and need to be tested.

Further complicating the team dynamic will be the new approach to the work environment whereby the traditional office setting has evolved from a brick-and-mortar building to an interconnected web of technology that sees many employees working remotely (Yawson, 2022). Developing teams in the remote world will require attention as well as novel interventions to achieve similar outcomes as previously experienced in current research. This paradigm shift is the result of the COVID pandemic. Though elements of current research may hold true, there will undoubtedly be a need to reevaluate the capability in years to come. Experiential team and leadership OD interventions like the ROPES course will be essential.

Conclusion

The current dialogic approach challenges organizations, and organizational development scholars and practitioners, to reimagine the field every day. The ROPES course is a way to invigorate teams to achieve greater harmony while also enhancing the cultural aspects of an organization that values learning and the development of its employees. A well-planned, thoughtful approach to the ROPES experience takes the learning from the course to the boardroom but can do so in more than just a cliché way. An organizational development practitioner can leverage the ROPES course as an effective tool for team building but this must be done methodically, with a clear understanding of an organizations strategic vision, and the participation of all team members throughout the intervention process.

References

Anderson, D. L. (2020). *Organization development the process of leading organizational change* (Fifth). SAGE.

Bjugstad, K., Thach, E. C., Thompson, K. J., & Morris, A. (2006). A fresh look at followership: A model for matching followership and leadership styles. *Journal of Behavioral and Applied Management*, *7*(3). https://doi.org/10.21818/001c.16673

Carsten, M. K., Uhl-Bien, M., West, B. J., Patera, J. L., & McGregor, R. (2010). Exploring Social Constructions of Followership: A qualitative study. The Leadership Quarterly, 21(3), 543–562. https://doi.org/10.1016/j.leaqua.2010.03.015

Chesley, J. A., Egan, T., & Jones, H. E. (2020). Elevating leadership development practices to meet emerging needs. *Journal of Leadership Education*, *19*(4), 180–189. https://doi.org/10.12806/v19/i4/t3

Coghlan, D. (1994). Managing organizational change through teams and groups. *Leadership & Organization Development Journal*, *15*(2), 18–23. https://doi.org/10.1108/01437739410055335

Daniels, M. R. (1994). Outdoor Adventure and Organizational Development: A ROPES Course Intervention. *Public Administration Quarterly*, *18*(2), 237–249. https://doi.org/https://www.jstor.org/stable/40861621

Daniels, M. R. (1997). The ROPES Course: Applications for Education and Training. *Public Administration Quarterly*, *3*(2), 239–242. http://www.jstor.org/stable/40215177

Eatough, E., Chang, C.-H., & Hall, N. (2015). Getting roped in: Group cohesion, Trust, and efficacy following a ROPES course intervention. *Performance Improvement Quarterly*, *28*(2), 65–89. https://doi.org/10.1002/piq.21183

Griffin, G., & Moors, R. (2004). The Fall and Rise of Organising in a Blue-Collar Union. *Journal of Industrial Relations*, *46*(1), 39–52. https://doi.org/10.1111/j.0022-1856.2004.00127.x

Haras, K., Bunting, C. J., & Witt, P. A. (2006). Meaningful involvement opportunities in ROPES course programs. *Journal of Leisure Research*, 38(3), 339–362. https://doi.org/10.1080/00222216.2006.11950082

Lacerenza, C. N., Marlow, S. L., Tannenbaum, S. I., & Salas, E. (2018). Team development interventions: Evidence-based approaches for improving teamwork. *American Psychologist, 73*(4), 517–531. https://doi.org/10.1037/amp0000295

Mandal, M. (2022). Team Building and Appreciative Inquiry in Research and Development Teams. *Vision*, 0(0). https://doi.org/10.1177/09722629211063528

Marczak, E. L., & Yawson, R. M. (2021). Understanding the Theories and Interventions of Motivation in Organization Development. *In, Responsible Management: Opportunities and Challenges. Proceedings of 58th Annual Conference of the Eastern Academy of Management.*, 1–22. https://doi.org/10.31124/advance.14610291.v1

Miller, D. J. (1998). *The impact of low ROPES course training on Team Development* (dissertation). UMI, Ann Arbor, MI.

O'Bannon, P. M. (2000). *The effects of the outdoor experiential ROPES course on perceived team performance* (dissertation). UMI, Ann Arbor, MI.

Osafo, E., & Yawson, R. M. (2021). Tempered radicalism as an approach to revisiting indigenous forms of critical human resource development. *European Journal of Training and Development*, *45*(2/3), 259–283. https://doi.org/10.1108/EJTD-03-2020-0049

Pearce, C. L., & Sims, H. P., Jr. (2002). Vertical versus shared leadership as predictors of the effectiveness of change management teams: An examination of aversive, directive, transactional, transformational, and empowering leader behaviors. *Group Dynamics: Theory, Research, and Practice, 6*(2), 172–197. https://doi.org/10.1037/1089-2699.6.2.172

Rohnke, K. (2010). *Silver bullets: A revised guide to initiative problems, adventure games, stunts, and Trust Activities*. Project Adventure.

Salas, E., Reyes, D. L., & McDaniel, S. H. (2018). The science of teamwork: Progress, reflections, and the road ahead. *American Psychologist, 73*(4), 593–600. https://doi.org/10.1037/amp0000334

Starr, M. (2004). *The experience of being a leader during a ROPES course program and at work: A heuristic inquiry* (dissertation). Boca Raton, FL.

Wagstaff, M. (n.d.). *A history of challenge courses*. from https://cdn.ymaws.com/www.acctinfo.org/resource/resmgr/imported/HistoryOfCCourses.pdf

Yawson, R. M. (2022). The future of work and becoming an employer of choice. *Organization Management Journal*, *19*(3), 86–87. https://doi.org/10.1108/OMJ-05-2022-970

Supply Chain Management and Logistics

Developing Supply Chain Resilience to Disruptions: Mindful Organizing and Supply Chain Relationships

ABSTRACT

The effectiveness of a supply chain in managing disruptive events depends on the organization's supply chain resilience, a critical element of modern management. Firms are increasingly exposed to supply chain disruptions that affect their operations. This research seeks to identify the key determinants of supply chain resilience among supply chain elements and types of relationships. It also specifically investigates how all these determinants vary in importance based on the tier of the extended supply chain the organization resides in. It investigates the effects of variables such as supply chain benevolence, risk management infrastructure, mindful organizing, top management orientation, disruption orientation, supply chain norms, and credibility on supply chain resilience. We collected data from global manufacturing firms using a survey, which was analyzed in a moderated-mediated model. The findings confirm the complexity of supply chain orientation via multiple interactions such as supply chain norms and credibility, which had a significant positive effect on supply chain resilience. Benevolence was also found to positively affect supply chain resilience between external manufacturers and their sub-tier supplier of components, but not in the external manufacturer-OEM relationship. The findings further suggest that firms must also be patient, defend their supply chain relationships, and be willing to make cooperative changes with supply chain partners to strengthen supply chain resilience. The findings also showed that mindful organizing can significantly influence resilience, but counterintuitively, can also pull time and energy away from more productive activity within the supply chain. This research finally discusses research implications and limitations.

Keywords: Supply Chain Resilience, Disruptions, External Manufacturers, Sub-tier Suppliers, Relationships

INTRODUCTION

Today, the global reach and dispersed nature of supply chain networks and respective lengths form an important context for the definition of the supply chain. Supply chains are dispersed and connected networks in which information and materials flow upstream and downstream. Resilience in these networks is ever critical in uncertain times. Sheffi and Rice (2005) [1] argue that while firms can build resilience in their supply chains, they do not assume the original state is achieved post-disruption, only that a stable state is re-established. The survival of a firm relies upon material being in the right place in the supply chain, at the right time, creating a competitive advantage for the end customer, hence the urgency to build a resilient supply chain (Sheffi & Rice, 2005) [1].

Prior researchers have focused on risk mitigation strategies to enable firms to develop resilient supply chains (Blackhurst, Dunn, & Craighead, 2011) [2]. Manuj and Mentzer (2008) [3] focus on supply chain flexibility, demand and supply, and risk management team contributions to resilience. Supply chain risk factors are well documented in the literature identifying risk associated with supply chain management (SCM) along with associated phases of disruptions (before, during, and post-disruption) and associated graphical profiles of disruption (Sheffi, 2007) [4]. Two risk categories are defined by Manuj and Mentzer (2008) [3] to be from within the supply chain or to be from the environment. A 2x2 matrix plotting probability of disruption versus consequences level is researched and defined by Sheffi (2007) [4]. Risk management strategies have been defined and discussed by Christopher and Peck (2004) [5], De Loach (2000) [6], and Manuj and Mentzer (2008) [3]. All agree risk management must be a core competency within firms as change and disruption are part of the business environment. Research by Craighead, Blackhurst, Rungtusanatham, and Handfield (2007) [7] noted that “the ability to share information visibly and

transparently – capacity to warn" looks beyond a risk mitigation list, a matrix, but understands that there needs to be a culture within a firm that enables resilience beyond tactical lists and risk mitigation actions. There is a gap in the literature relative to supply chain social relationships and interconnectedness effect on supply chain resilience during global disruptions. This research fills that gap and provides the first comprehensive look at the role that key elements in the supply chain affect its resilience.

MANAGING DISRUPTIONS

Supply chain disruptions are events that are characterized by "*high uncertainty*" (Bode, Wagner, Petersen, and Ellram, 2011) [8] and "*disrupt the normal flow of goods and services within the supply chain"* (Craighead et al., 2007) [7]. The high uncertainty in the globally dispersed supply chain and firms responding to disruptions creates the need for a mindset that looks ahead and anticipates disruption to build a resilient node and interconnected dyads within the supply chain. Each disruption over time adds knowledge, lessons, and learning among the nodes which enhances a mindset (orientation) towards thinking about disruptions. This enables continuous updating of institutional knowledge and actions by firms within the network to adapt and learn, scaling up the level of resilience post-disruption. It is this supply chain disruption orientation (SDRO) that provides a footing for resilience. Each disruption experience in the supply chain adds to a firm's disruption orientation, adding to its supply chain risk management program (SCRM) (Ambulkar, Blackhurst, & Grawe, 2015) [9]. If the firms in the supply chain are disposed towards operating with a mindset geared towards high uncertainty, that leads to higher resilience in the supply chain (Ambulkar et al., 2015) [9].

The research of Bode et al. (2011) [8] leverages market-orientation literature to define supply chain disruption orientation as *"a firm's general awareness and consciousness of, concerns about,*

seriousness toward, and recognition of opportunity to learn from supply chain disruptions". The link to stability is not only parallel to the discussion by Holling (1973) [10], relative to our ecological system, but also touched upon by Bode et al. (2011) [8], who notes that "*the stronger the firm's supply chain disruption, the more importance it attaches to the issue of supply chain disruptions and the more pronounced its need for stability is*" (Bode et al., 2011) [8]. Strong SDRO drives a firm to act when there are disruptions and these firms are known as "active firms" (Daft & Weick, 1984) [11] that learn from disruptions.

The SDRO construct was developed by Bode et al. (2011) [8] in a model to define organizational responses to supply chain disruptions. The model by Bode et al. (2011) [8] related SDRO specifically to a buffering and bridging organizational response. Ambulkar et al. (2015) [9] take Bode et al. (2011) [8] SDRO construct further, including it as an antecedent to firms developing resilience to supply chain disruptions combined with a resource management construct. The origination of SDRO with Bode et al. (2011) [8] research study includes the organizational response to adverse environmental events at the dyad level. Ambulkar et al. (2015) [9] model determines SDRO is not the only factor affecting supply chain resilience, and added a resource reconfiguration factor and a risk management infrastructure (RMI) factor, measuring the effect on supply chain resilience (SCR). Ambulkar et al. (2015) [9] defines the resilience construct for use in this research model as the dependent variable, supply chain resilience (SCR).

FRAMEWORK

Although existing research relative to the supply chain disruption orientation (SDRO) construct has been associated with an increase in supply chain resilience, this research includes the addition of antecedents contributing to the level of supply chain resilience beyond SDRO, such as supply chain orientation (SCO), key supplier relationship management (KSRM), risk management

infrastructure (RMI), combined with mindful organizing (MO) as factors affecting the level of supply chain resilience (SCR) during disruptive events. The five independent variables are existing constructs with verified scales, but have not been assessed as mediators or moderators when combined with SDRO and SCR. The contribution to the supply chain professional network includes research into the complex supplier relationship factors with expectations identifying those factors that increase supply chain resilience. There is also no prior research on the combined benefits of mindful organizing (MO) with other contextual factors that potentially enable supply chain resilience.

SCO and KSRM have been studied together relative to organizational buying effectiveness (Miocevic & Crnjak-Karanovic, 2012) [12] , but not supply chain resilience. Resilience to supply chain disruptions has been researched where a firm resilience (SCR) scale was developed (Ambulkar et al., 2015) [9] and used alongside an SDRO scale developed by Bode et al. (2011) [8]. The risk management infrastructure scale (RMI) was developed (Ambulkar et al., 2015) [9] as an antecedent to the firm's resilience to supply chain disruptions, while using the construct of resources configuration as a mediator. That prior research added to the literature an additional definition for resilience at the firm level. This new research will leverage the SDRO scale from Bode et al. (2011) [8] and RMI and SCR scales from Ambulkar et al. (2015) [9]. These antecedents of resilience have not been comprehensively researched as combined antecedents relative to SCR.

HYPOTHESES

Supply Chain Disruption Orientation (SDRO). Supply chain disruptions are events that are characterized by "high uncertainty" (Bode et al., 2011) [8] and "disrupt the normal flow of goods and services within the supply chain" (Craighead et al., 2007) [7]. The high uncertainty in the globally dispersed supply chain and firms responding to disruptions creates the need for a mindset

that looks ahead and anticipates disruption to build a resilient node or dyad within the supply chain. A mindset of thinking about disruptions and updating the firms within the network to adapt and learn is SDRO that provides a foundation for resilience.

Hypothesis 1: The greater the supply chain disruption orientation (SDRO), the greater the supply chain resilience (SCR).

Supply Chain Orientation (SCO). Mentzer, Dewitt, Keebler, Min, Nix, Smith & Zacharia (2001) [13] noted that SCO inside a firm is required to enable supply chain management across firms in the supply chain. The six constructs within SCO include the concept of norms for setting relationships defining how to act and behave within the supply chain. Included in SCO is a measure to determine if the supply chain members' actions and behaviors are congruent with a protective orientation that includes looking out for each other within the network of the supply chain. The more transparency and sharing across firms in the supply chain, then the more it enhances the SDRO for each firm. SDRO and resilience is positively increased when mediated by SCO. The six elements of SCO include credibility, benevolence, commitment, norms, compatibility, and top management support (Min & Mentzer, 2004) [14]. These six elements together form SCO that leads to resilience as indicators to the level at which supply chain firms work together. The link between SDRO and SCR is better explained by the mediating role of SCO. SCO adds positively to the relationship between SDRO and resilience.

Hypothesis 2: The effect of supply disruption orientation (SDRO) on supply chain resilience (SCR) is partially mediated by supply chain orientation (SCO).

Ambulkar et al. (2015) [9] prior research confirms that firms in the supply chain that have a disposition for operating with a mindset for high uncertainty (SDRO) leads to resilience in the supply chain (Ambulkar et al., 2015) [9] (SCR). This study posits that there is an intervening variable, SCO. SDRO leads to SCO, which in turn leads to SCR. The SCO construct, defined as one variable, combines orientation for credibility, benevolence, commitment, norms,

compatibility, and top management support (Min & Mentzer, 2004) [14]. The firms with heightened awareness of disruptions (SDRO) leads to overall supply chain orientation (SCO), which in turn leads to supply chain resilience (SCR). Hence, SDRO has an indirect effect on SCR through SCO.

Key Supplier Relationship Management (KSRM). "Key supplier relationship management (KSRM) focuses on the management of strategic relationships and is based on the assumption that the firm's overall portfolio comprises relationships that may have different levels of importance" (Miocevic & Crnjak-Karanovic, 2012) [12]. This is a critical distinction that a transactional relationship is not really one with a partner, versus a key relationship which pushes a deeper partnership and interconnected relationship with more expectations among the nodes within a supply chain. There is prior literature linking SCO and KSRM concepts to firm outcomes (Min & Mentzer, 2004) [14] and purchasing behavior (Miocevic, & Crnjak-Karanovic, 2012) [12]. This present study argues that KSRM partially mediates between SDRO and SCR. The intervening variable, KSRM, has two phases such that the early phase effect will strengthen the effect of SDRO on SCR. The later phase of KSRM will weaken the effect of SDRO on SCR. This partial mediation, when KSRM exists, can weaken SDRO on SCR as the firm is not truly protected as imagined during a disruption to the supply chain. Supply chain terms and conditions may be strong at the outset and vested between dyads, but as time elapses and changes in the supply chain, like disruptions, the belief of protection from impacts is false. SDRO keeps pace being alert and being aware, but the KSRM establish written agreements that hold for lengths of time, after legal agreements are executed between firms. KSRM agreements become uncoordinated and are not able to change in real-time. When the firm experiences an external environment change, KSRM agreements are not typically amended during these emergent events. Therefore, the effect of SDRO

on SCR during the early phases through KSRM is partially mediated (stronger) but then the effect of SDRO on SCR partially mediates (weakens) when KSRM is further along in its relationship (control phase of KSRM).

> *Hypothesis 3a: The effect of supply disruption orientation (SDRO) on supply chain resilience (SCR) is partially mediated by key supplier relationships (KSRM), during the planning and implementation phases of KSRM.*
> *Hypothesis 3b: The effect of supply disruption orientation (SDRO) on supply chain resilience (SCR) is partially mediated by key supplier relationships (KSRM), during the control phases (post planning and implementing) of KSRM.*

Mindful Organizing (MO). There are five elements that define the mindful organizing (MO) concept that originated with Vogus and Sutcliffe (2007, 2012) [15][16]: preoccupation with failure, reluctance to simplify interpretations, sensitivity to operations, commitment to resilience, and deference to expertise. It is within front line nursing units where these constructs were first applied and subsequently contribute to the theory linking to High Reliability Organizations (HRO). The construction and validation of the Mindfulness Organizing Scale (MOS) originated originally as the Safety Organizing Scale (SOS) applicable to front line nursing units to prevent mistakes in these units.

In terms of MO applicability to supply chain resilience, the capabilities that MO includes are sensing, seizing, and reconfiguration. Resilience is positively impacted by supply chains that can reconfigure, but importantly, also respond and manage a supply chain via sensing. Supply chains create a flurry of questions and urgency when emergent events occur, and MO assesses these processes that add to resilience. There is increasing qualitative and quantitative evidence that mindful organizing creates and sustains the superior performance of HROs (Vogus & Sutcliffe, 2007) [15]. Early detection of emergent events and adapting is mindful organizing. MO depends upon sensing and uncovering concerns, which enables action more readily and improves resilience during unexpected events (Berente, Lyytinen, Yoo, & King, 2016) [17]. The ability to detect

issues early and internalize within supply chain firms increases the level of resilience. Mindful organizing actions in organizations moderate the level of resilience within the supply chain. Higher mindful organizing (MO) will strengthen the relationship between supply chain disruption orientation (SDRO) and supply chain resilience (SCR).

Hypothesis 4a: Increases in MO will strengthen the relationship between SDRO and SCR.
Hypothesis 4b: Increases in MO will strengthen the relationship between SDRO and SCO.
Hypothesis 4c: Increases in MO will strengthen the relationship between SCO and SCR.
Hypothesis 4d: Increases in MO will strengthen the relationship between SDRO and KSRM.
Hypothesis 4e: Increases in MO will strengthen the relationship between KSRM and SCR.

Risk Management Infrastructure (RMI). Ambulkar et al. (2015) [9] defines RMI as "a firm's structure of resources designed to manage risk in the supply chain" (Ambulkar et al., 2015) [9]. Three attributes defined by Ambulkar et al. (2015) [9] are included in RMI, namely information systems, use of key performance indicators (KPI) and metrics, and the investment in personnel to manage supply chain risks and disruptions. RMI focus is on the organizational capacity attention to disruptions within a firm. When organizations increase RMI investments, increases in SCR are expected. The investments in both personnel capacity and information technology capability enable supply chain risk management (SCRM). Increasing SCRM creates readiness for novel events and metrics provide insights relative to specific events. Also, information systems enhance resilience to manage supply chain risks and disruptions, thus increasing resilience when there are emergent events. The more investment in risk management resources, the greater the increase in resilience.

Hypothesis 5a: Increases in risk management infrastructure (RMI) will strengthen the relationship between SDRO and SCR.
Hypothesis 5b: Increases in risk management infrastructure (RMI) will strengthen the relationship between SDRO and SCO.
Hypothesis 5c: Increases in risk management infrastructure (RMI) will strengthen the relationship between SCO and SCR.
Hypothesis 5d: Increases in risk management infrastructure (RMI) will strengthen the relationship between SDRO and KSRM.

Hypothesis 5e: Increases in risk management infrastructure (RMI) will strengthen the relationship between KSRM and SCR.

Ambulkar et al. (2015) [9] argue that risk management infrastructure (RMI) does not have the same benefit when there are "high impact disruptions". The argument in the literature is that the rigid processes, metrics, information systems, and approaches apply to routine everyday disruptions, but do not allow the same benefits when an organization is faced with a "high impact disruption".

METHODOLOGY

Data and Procedure. To test the hypotheses on an empirical basis, this research conducted a survey between May 2021 and July 2021 using a Qualtrics internet-based survey provided to supply chain professionals via email. The email was sent out directly to contacts and network of individuals that had any contact with suppliers in their roles. These supply chain contacts were colleagues of the researcher, and each was asked to forward to their contacts that worked in supply chain functions such as procurement, engineering, execution, materials, planning, and operations. The researcher personally sent out the original request to potential participants a request for each to forward the email (and link) request to their colleagues and sub-tier suppliers from within their network. The Qualtrics Survey link was cascaded to professionals working at OEMs, EMs, and layers of sub-tier suppliers, inclusive of sub-tier suppliers such as component manufacturers and raw material sources. The survey did not target specific roles but diverse respondents, including business and engineering professionals who interact within supply chain relationships regardless of tier or level within the supply chain

Primary data. There were 313 respondents, comprising 172 OEMs, 76 External Manufacturers, and 43 drawn from various sub-tier levels in the supply chain. Additionally, there were 22 respondents from the logistics arm of the supply chain. The survey had a global reach with 69.23%

of the respondents from North America (US, Canada, and Mexico), 21.79% from Asia, 4.80% from Europe, and 4.17% from other countries. The respondent's role within their firms included 20.32% in the Supply Chain Management function, 15.81% Program Managers, 9.35% Design Engineers, 9.35% Operations Engineers, 8.06% Demand Planners, 7.42% Quality Engineers, 6.45% Purchasing, 5.81% Logistics, and 17% other functions. Most of the respondents (76.22%) were from the high-tech sector, followed by 9.45% in the industrial sector, 4.23% from the consumer sector, 1.30% from the automotive and medical sectors, and the remaining 8.79% from other market segments. The data collection contacts were heavily weighted to the high- technology sector. The demographic statistics of the respondents is detailed in Table 1: A. number (N) of respondents, B. respondent level/tier, C. respondent geography, D. respondent role/function, and E. respondent industry.

	Tier	Location	Function	Industry
Number(N) Respondents	313	312	310	307

Table 1.A: Respondents: Number

Supplier Tier	N	Percentage
1= OEM	172	54.95%
2= External Manufacturer	76	24.28%
3= Sub-tier Assembly	24	7.67%
4= Sub-tier Component	14	4.47%
5= Raw Material	5	1.60%
7 = Logistics & Other	22	7.03%

Table 1.B: Respondents: Tier

Firm Location	N	Percentage
1= USA, Canada, Mexico	216	69.23%
2= Asia	68	21.79%
3= Europe	15	4.81%
7= Other	13	4.17%

Table 1.C: Respondents: Geography

Function/Role	N	Percentage
Supplier Management	63	20.32%
Program Management	49	15.81%
Design Engineering	29	9.35%
Operations Engineering	29	9.35%

Planning	25	8.06%
Quality	23	7.42%
Purchasing	20	6.45%
Logistics	18	5.81%
Other	54	17.42%

Table 1.D: Respondents: Role/Function

Manufacturer's Industry	N	Percentage
High-Technology (e.g., PCs, laptops, servers, printed circuit boards, memory, hard drives)	234	76.22%
Industrial Products (e.g., capital goods, raw materials, component parts, major equipment, operating suppliers, and services)	29	9.45%
Consumer Products (e.g., convenience products, shopping products, specialty products)	13	4.23%
Medical/Health Products (e.g., devices, home care, hospital care)	2	0.65%
Automotive Products (e.g., parts, vehicles)	2	0.65%
Other	27	8.79%

Table 1.E: Respondents: Industry

Survey Instrument. A survey methodology was used to collect the data and test the proposed hypotheses during 2021. The survey was developed based on literature review and existing constructs. The review considered empirical studies in the fields of supply chain management, purchasing and operations, marketing, and organization science. The survey instrument comprised 52 items for the respondents to consider relative to their own supply chain relationships. All survey items were derived from existing literature (see Table 12: Measurement Scales). Pre-established scales were used to measure supply chain disruption orientation (SDRO), supply chain orientation (SCO), key supplier relationship management (KSRM), risk management infrastructure (RMI), mindful organizing (MO), and supply chain resilience (SCR).

The survey instruments did not provide details to the respondents of the actual relationships under investigation. The survey provided anonymity and confidentiality to the respondents. This was used to reduce the chances of responses being swayed to a desired response perceived by responders for the research. Respondents were not asked to base their answers on any specific supply chain disruption even though the timing of the survey was one year into the Covid-19

supply chain disruption. In early January 2020, the impacts were not as visible within the tiers (layers) of the supply chain but as the year progressed, country and factory closures increased, global death rates rose and so did the supply chain disruptions. One and a half years later, supply chain relationships were characterized by disruptions and the economy was far from recovery to pre-pandemic lead-time availability of goods and services. The survey was administered within and across various supply chain tiers (levels).

Measures. The research operationalized constructs according to the existing literature in a novel combination of antecedents relative to supply chain orientations, disruption orientations, risk management infrastructures, key supplier relationships, and mindful organizing. These were postulated to influence supply chain resilience. This was also investigated in the context of the dyadic relationships between adjacent tier levels within the supply chain. The constructs were included as antecedents to supply chain resilience. The model including the dependent variable, independent variables, and controls are depicted in Figure 1: Hypotheses. This figure includes the complexity of two parallel mediators and two moderators.

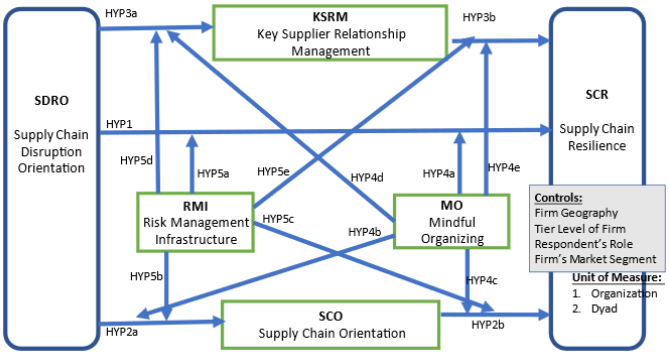


Figure 1: Hypotheses

Dependent variable. The Supply Chain Resilience (SCR) construct was operationalized using the established four items: the ability of the firm to cope with changes due to a supply chain disruption, the ability to adapt to a supply chain disruption, provide a quick response, and the ability to maintain high

situational awareness. These constructs were adopted from Ambulkar et al. (2015) [9]. The four items were measured on a seven-point Likert scale (1 = strongly disagree, 7 = strongly agree).

Independent variables. The Supply Chain Disruption Orientation (SDRO) construct was adopted from Bode et al. (2011) [8]. This supply chain disruption orientation measured a firm's response to disruption specifically focused on taking actions relative to buffering or bridging a firm's resources as responses to supply chain disruptions. This construct was introduced as supply disruption orientation. The SDRO construct includes the ability to learn from supply chain disruptions along with a culture of persistent alertness and awareness to supply chain disruptions, such that a firm is sensitized and conscious that disruptions are imminent. Bode et al. (2011) [8] constructs target behaviors applied to this research focused on resilience as a response to supply chain disruptions. The scale consists of five items measured on a seven-point Likert scale (1=strongly disagree, 7-strongly agree).

Supply Chain Orientation (SCO) was measured using a multidimensional scale adopted from Min and Mentzer (2004) [14]. This scale focused on relational aspects of supply chain management. The SCO construct was based on the implications of managing supply chain flows (Mentzer et al, 2001) [13]. There are five dimensions with a total of 19 items. SCO structural constructs focus on behavioral elements that facilitate relational exchange across supply chain nodes. These behavior constructs for SCO include the following cultural dimensions defined as: credibility, benevolence, commitment, norms, organizational compatibility, and top management support. The mindset within supply chain-oriented firms facilitates trust, commitment, compatible supply chain partners to cooperate with, and top management endorsement (Esper, Deffee, & Mentzer, 2010) [18]. Min and Mentzer (2004) [14] concluded that supply chain-oriented firms should nurture, maintain, and improve these five elements inside the firm when working with its supply chain relationships (Min & Mentzer, 2004) [14].

The Key Supplier Relationship Management (KSRM) construct was operationalized based on the work of Miocevic and Crnjak-Karanovic (2012) [12]. It has three dimensions: planning, implementation, and control. Supplier relationships are not all the same, as there are key relationships aligned on strategic direction. There are also key partners for both the supplier and customers, beyond the buying and selling transactions required for everyday supply chain activities (Miocervic & Crnjak-Karanovic, 2012) [12]. These KSRM constructs measure three phases of key relationships: planning, implementation, and control phases.

The Risk Management Infrastructure (RMI) construct was operationalized by Ambulkar et al. (2015) [9]. Blackhurst et al. (2011) [2] contended that investments in organizational resources provide assets, measure, and monitors that increase a firm's supply chain resilience when there are supply chain disruptions. Investments in RMI are measured by Ambulkar et al. (2015) [9] using three items: (a) people assigned to manage supply chain risks and/or disruptions, (b) metrics (key performance indicators) monitoring supply chain risk/disruptions, and (c) information systems in place to manage supply chain risks and disruptions.

The Mindful Organizing (MO) construct was developed by Weick and Sutcliffe (2006) [19], based on mindful organizations being those defined to spend more time on five aspects: (a) examining failure looking into the "health of the system", (b) resisting the urge to simplify assumptions about "the world", (c) observing operations and their effects, (d) developing resilience to manage unexpected events, and (e) locating local expertise and creating a "climate of deference" to those experts (Vogus & Sutcliffe, 2007, 2012) [15,16]. The MO is an eight-item construct with dimensions consisting of five interrelated behavioral processes that form the MO scale: preoccupation with failure, reluctance to simplify interpretations, sensitivity to operations, commitment to resilience, and deference to expertise (Vogus & Sutcliffe, 2007, 2012) [15,16].

Vogus and Sutcliffe (2007, 2012) [15,16] constructed and validated the Mindfulness (formerly Safety) Organizing Scale.

Control variables. This study included four control variables that could affect supply chain resilience (SCR): respondent's firm industry, firm location (geography), respondent's job functional role, and the firm's tier-level of operations within the supply chain. This included OEMs and first tier to n^{th} tier firms that produce and deliver goods and services across the supply chain. The supplier's industry is a control as a firm's market segment may be exposed to different disruptions than others, affecting supply chain resilience differently between manufacturing industries. The supplier's global location may affect supply chain resilience if some geographic regions experience disruptions that vary in frequency, and internal or internal disruptions, thus resilience may differ. Behaviors and cultures in regions may also differ when measuring the effect on resilience. The tier level that a supplier provides goods and services within a supply chain is noted as the supplier's tier within the supply chain.

Depending upon the tier within the supply chain engagement, each may be exposed to disruptions at varying times, magnitudes, and frequencies. For example, the OEM visibility to disruptions at the raw material source level (tier) may be lagged in timing and effects and visibility compared to the tier that is directly responsible for the raw material. The suppliers that convert raw materials into parts and components are several tiers removed from the disruption and the OEM is removed even further from the origin of the disruption and this difference in tier-level in the supply chain may affect differences in supply chain resilience.

The last of the four controls is the respondent's job (functional role) within the supply chain that each represents. For example, a planner, a program manager, or engineer each may engage in different types of disruption that touch their span of control that is specific to their function. Roles

and responsibilities may segment responses to disruptions, such as responses to stockouts and lead-time extensions vs. quality yield losses and manufacturing process disruptions touching each functional role at different times. All control variable constructs and responding scale items can be found in Table 12: Measurement Scales.

Measurement Assessment. Prior to measurement assessment, surveys with missing values were removed from the data set. The data set was reduced to 304 respondents who filled out all independent and dependent variable items. The data set was further reduced to 296 respondents who entered all the four control variables and answered questions for all the independent and dependent variable constructs.

Confirmatory factor analysis. Reliability, validity, and dimensionality of the 52 items were assessed using a confirmatory factor analysis (CFA). The psychometric properties by means of a covariance-based CFA was completed. All independent and dependent variables were included in a single multifactorial CFA model. The measurement model appears in Table 12: Measurement Scales and includes the inter-construct correlations and average variances extracted.

First-order measurement model. Although this research deployed existing constructs in the survey, all these leveraged constructs were used in a novel combination to measure the relationship to supply chain resilience. The reliability and validity analysis were completed on the 52 leveraged scales, as 6 factors. Exploratory factor analysis was conducted on all the independent and dependent variable constructs in SPSS using principal component analysis, varimax rotation, and extraction criterion of eigenvalue greater than 1.00. The factor analysis revealed that one element of the SCO was reverse coded (SCO_CRED_4). This one response was recoded (SCO_CRED_4) and CFA was repeated.

Analysis of the existing 52 scales provided a nine-factor solution (originally 6 factor) that explained 67.580% of the variance. The resulting Keyser-Meyer-Olkin (KMO) measure of sampling adequacy was .939 and the Chi-Square was 9395.612 (df. = 1035, p=0.000). The KMO measure of 0.939 (greater than recommended value of 0.6) suggest that the sample is adequate for factor analysis and that the factor analysis is likely to provide reliable factors. The significance of the Chi-Square value also suggests that a nine-factor analysis is significant. Convergent validity and reliability were assessed using factor loadings and Cronbach's Alpha values, respectively. All retained indicators had factor loadings higher than 0.500 ($p<0.001$) and Cronbach Alpha >0.600, per Table 12: Measurement Scales. The reliability level of the instrument is represented by the Cronbach Alpha. A Cronbach Alpha value above 0.600 is considered high reliability and acceptable index (Nunnally & Bernstein, 1994). Values in the range of 0.60 - 0.80 are considered moderate, but acceptable per the literature.

Formative second-order measures. The leveraged constructs from literature (SDRO, SCO, KSRM, RMI, MO, and SCR) had multiple items within each that established each of the respective final formative constructs. The factor loadings maintained only two of the four original SCO multi-dimensional factors from SCO (SCO_TopMgt, SCO_BEN) while retaining four of the other original constructs and the respective original items intact included within the hypotheses, namely: KSRM, MO, SCR, RMI. There were three new composite dimensions after dropping one dimension that failed the reliability test (SCO_CRED_4). The original SCO_CRED dimension reduced from 4 elements to 2 elements (SCO_CRED_2 and SCO_CRED_3). The SDRO dimension reduced from 5 elements to three plus added one item from the SCO_COMM_NORM (SDRO_1, SDRO_2, SDRO_3, SCO_COMM_NORM_4). The SCO dimension for Commitment and Norms reduced from the original seven items to final PCA loading with three items

(SCO_COMM_NORM_1, SCO_COMM_NORM_2, SCO_COMM_NORM_3). The original SCO items did not hold together as one construct and required the final analysis to use SCO in its new form of 4 constructs and not one consolidated construct as defined in the hypotheses.

The final nine composite constructs based on the PCA, reliability, and validity test were coded using the average of the individual items loading per PCA. The natural log for each final formative composite construct was used in the model based on the normality graphs/plots. Factor loading, and reliability are shown in Table 12: Measurement Scales, respectively. The original one measure of SCO was broken apart as the loadings did not keep SCO together as one measure. The analysis was based on the loadings as such the original model with six constructs changed to nine constructs with SCO broken into different elements than originally leveraged.

Dependent Construct. The dependent variable supply chain resilience (SCR) and its items remained intact based on PCA loadings all >.500 and re-labeled SCRALL included SCR_1, SCR_2, SCR_3, and SCR_4 (no change from the leveraged construct). The composite Cronbach Alpha was 0.880 for this construct. The analysis model used the natural log for this construct, which was labeled Res_ln.

Independent Constructs. Three of the original independent variable composite constructs and their respective items for each remained intact based on PCA loadings (>0.500) for key supplier relationship management (KSRM -12 items), mindful organizing (MO - 8 items), and risk management infrastructure (RMI – 3 items). These three constructs were re-labeled as KeySup, Mind, and RiskMgt with Cronbach Alphas for the formative constructs of 0.935, 0.915, 0.836 respectively (analysis model included natural logs labeled Key_ln, Min_ln, and RSK_ln respectively).

The factor loadings maintained two of the four original SCO multi-dimensional factors from SCO composite (SCO_TopMgt, SCO_BEN) labeled in the new nine-construct model as SCOMgt and SCOBEN with Cronbach Alphas for the formative constructs of .901 and .874 respectively (natural logs for the analysis model were labeled Mgt_ln and Ben_ln respectively).

Three new independent formative constructs were developed based upon PCA, which resulted in dropping six of the original items with loadings <0.500 per Table 12: Measurement Scales. DISRUPT construct added an item based on PCA loading and dropped 2 items based on loadings less than 0.500 resulting in the final construct labeled DIS_ln. The construct SCONorm dropped three items as PCA loadings for items less than 0.500 (SCO_COMM_and_Norm items 5, 6, and 7) were dropped and not within the final construct labeled Norm_ln. The SCOCRED dropped one item for low PCA loading (SCO_CRED_1),0.461, then relabeled the new final construct as SCOCR_ln. The final 45 items from the original 52 items resulted in a nine-factor (constructs) from the original six constructs, see Table 12: Measurement Scales. Table 2 explains the total variance for the new nine components.

Total Variance Explained									
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	17.724	38.531	38.531	17.724	38.531	38.531	6.917	15.037	15.037
2	2.153	4.680	43.211	2.153	4.680	43.211	5.481	11.914	26.951
3	2.045	4.447	47.658	2.045	4.447	47.658	3.684	8.009	34.960
4	2.010	4.369	52.026	2.010	4.369	52.026	3.143	6.832	41.792
5	1.862	4.048	56.075	1.862	4.048	56.075	3.033	6.593	48.385
6	1.561	3.394	59.468	1.561	3.394	59.468	2.624	5.704	54.089
7	1.432	3.112	62.581	1.432	3.112	62.581	2.244	4.879	58.968
8	1.206	2.622	65.203	1.206	2.622	65.203	2.094	4.553	63.521
9	1.093	2.377	**67.580**	1.093	2.377	**67.580**	1.867	4.059	**67.580**

Table 2: Total Variance Explained – Nine Components

Construct	Qty of Items	Cronbach's Alpha * Reliability	(N) Number of Respondents	Summed Items	VIF (predictors)	Construct
SCOCR	2	0.710	304	SCO_CRED_2+SCO_CRED_3 (removed CRED4_R alpha was .547 all 3))	***1.493***	SCOCRED
Key	12	0.935	304	KSRM_12+KSRM_11+KSRM_10+KSRM_9+KSRM_8+KSRM_7+KSRM_6+KSRM_5+KSRM_	***2.663***	KeySup
Min	8	0.915	304	MO_1+MO_2+MO_3+MO_4+MO_5+MO_6+MO_7+MO_8	***2.175***	Mind
Mgt	5	0.901	304	SCO_Top_MGT_1+SCO_Top_MGT_2+SCO_Top_MGT_3+SCO_Top_MGT_4+SCO_Top_M	***2.047***	SCOMgt
Res	4	0.880	304	SCR_1+SCR_2+SCR_3+SCR_4	Dependent Variable	SCRALL
Ben	4	0.874	304	SCO_BEN_1+SCO_BEN_2+SCO_BEN_3+SCO_BEN_4	***1.925***	SCOBEN
RSK	3	0.836	304	RMI_1+RMI_2+RMI_3	***1.709***	RskInfr
DIS	4	0.694	304	SDRO_2+SDRO_3+SDRO_1+SCO_COMM_and_NORM_4	***1.391***	DISRUPT
Norm	3	0.773	304	SCO_COMM_and_NORM_2+SCO_COMM_and_NORM_3+SCO_COMM_and_NORM_1	***1.725***	SCONorm

*Cronbach's Alpha > .700 high reliability; 0.60-0.80 moderate, but acceptable

Table 3: Reliability Statistics for Final Constructs

Validity of formative constructs. To determine the validity of the nine formative constructs we examined the variance inflation factors (VIFs) of the independent indicators. The regression of any of the independent variables on all the other independents (total of 8), ignoring the dependent variable, resulted in all VIFs less than 2.70 which indicates there is no problem with multicollinearity between the nine variables. Therefore, we summed the first-order constructs as follows creating the formative constructs in SPSS based on N=304 respondents, see Table 3: Reliability Statistics Final Constructs.

First order constructs summed (and averaged) then the natural log was operationalized for the model. Constructs defined as follows:

SCOCRED=(SCO_CRED_2+SCO_CRED_3)/2

KeySup=(KSRM_12+KSRM_11+KSRM_10+KSRM_9+KSRM_8+KSRM_7+KSRM_6+KSRM_5+KSRM_4+KSRM_3+KSRM_2+KSRM_1)/12

Mind=(MO_1+MO_2+MO_3+MO_4+MO_5+MO_6+MO_7+MO_8)/8

SCOMgt=(SCO_Top_MGT_1+SCO_Top_MGT_2+SCO_Top_MGT_3+SCO_Top_MGT_4+SCO_Top_MGT_5)/5

SCRALL=(SCR_1+SCR_2+SCR_3+SCR_4)/4

SCOBEN=(SCO_BEN_1+SCO_BEN_2+SCO_BEN_3+SCO_BEN_4)/4

RskInfr=(RMI_1+RMI_2+RMI_3)/3

DISRUPT=(SDRO_2+SDRO_3+SDRO_1+SCO_COMM_and_NORM_4)/4

SCONorm=(SCO_COMM_and_NORM_2+SCO_COMM_and_NORM_3+SCO_COMM_and_NORM_1)/3

Reliability, validity, and dimensionality of constructs were assessed using PCA. Factor loadings, composite reliability (CR), and average variance extracted (AVE) estimates were deployed to ensure convergent validity of the constructs. Factor loadings were not all >0.500 resulting in the final nine constructs being composed of 45 items, reduced from the original 52 items (six constructs) leveraged from prior supply chain literature research.

To achieve convergent validity, the average variance extracted (AVE) must exceed a threshold of 0.50 for every construct or the CR must exceed the threshold of 0.70. There were 4 constructs that do not have an AVE above the critical cut-off value of 0.50. But these four did have CR above 0.70. One of the constructs misses the Cronbach Alpha minimum 0.70 requirement recording at 0.694, but the CR for this construct is 0.752 meeting acceptable convergent validity. All the constructs had a CR >0.70 threshold, confirming the existence of convergent validity in this research because all the indicators measured the latent construct they are intended to measure. To examine discriminant validity, the squared correlations were reviewed between every pair of first order constructs. The highest correlation between two constructs must not exceed each construct's square root AVEs, and this criterion was accomplished (see Table 12: Measurement Scales), providing evidence of discriminant validity. The final nine first-ordered constructs yielded dimensionality, reliability, and validity. The VIF predictor were all less than 2.7, see Table 4: Correlation of Final Constructs. The higher the VIF value, the greater the correlation of the variable with other variables. Values of more than 4 or 5 might be regarded as being moderate to high, with values of 10 or more being regarded as very high. Our model is well within acceptable levels per Table 3, Reliability Statistics for Final Constructs.

The final 45 items were run in an additional exploratory factor analysis in order to inspect for possible cross-loadings. The KMO and Bartlett's Test on 45 items provided .940 KMO (df. =990, p< 0.000). The total variance explained is 68.366% with 45 items and nine constructs for the final model (increase of .786% total variance explained when compared to the original 52 items, six consolidated constructs). No cross-loadings were present in the data, per Table 4, as none of the factors share the same variable. Therefore, all final nine constructs are independent as the items load to single constructs, see Table 5: Discriminant validity of final nine constructs (average variance extracted and correlations).

Correlations										
		SCRALL	SCOCR	KeySup	Mind	SCOMgt	SCOBEN	RSKInfr	DISRUPT	SCONorm
Pearson Correlation	SCRALL	1.000	0.400	0.522	0.464	0.453	0.465	0.445	0.375	0.435
	SCOCR	0.400	1.000	0.481	0.475	0.404	0.396	0.393	0.348	0.464
	KeySup	0.522	0.481	1.000	0.667	0.626	0.633	0.564	0.460	0.510
	Mind	0.464	0.475	0.667	1.000	0.568	0.591	0.489	0.402	0.492
	SCOMgt	0.453	0.404	0.626	0.568	1.000	0.493	0.565	0.391	0.524
	SCOBEN	0.465	0.396	0.633	0.591	0.493	1.000	0.439	0.405	0.494
	RSKInfr	0.445	0.393	0.564	0.489	0.565	0.439	1.000	0.268	0.438
	DISRUPT	0.375	0.348	0.460	0.402	0.391	0.405	0.268	1.000	0.415
	SCONorm	0.435	0.464	0.510	0.492	0.524	0.494	0.438	0.415	1.000
Sig. (1-tailed)	SCRALL		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	SCOCR	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000
	KeySup	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000
	Mind	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000
	SCOMgt	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000
	SCOBEN	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000
	RSKInfr	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000
	DISRUPT	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000
	SCONorm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	

Table 4: Correlations of final nine constructs

	Mean	SD	SCRALL	SCOCR	KeySup	Mind	SCOMgt	SCOBEN	RSKInfr	DISRUPT	SCONorm
SCRALL	2.917	1.219	**0.773**								
SCOCR	2.288	1.003	0.400	**0.748**							
KeySup	2.455	0.917	0.522	0.481	**0.650**						
Mind	2.648	1.022	0.464	0.475	0.667	**0.696**					
SCOMgt	3.082	1.317	0.453	0.404	0.626	0.568	**0.713**				
SCOBEN	2.827	1.075	0.465	0.396	0.633	0.591	0.493	**0.716**			
RSKInfr	2.839	1.268	0.445	0.393	0.564	0.489	0.565	0.439	**0.720**		
DISRUPT	1.835	0.725	0.375	0.348	0.460	0.402	0.391	0.405	0.268	**0.657**	
SCONorm	2.714	0.994	0.435	0.464	0.510	0.492	0.524	0.494	0.438	0.415	**0.683**
Note: Bold numbers on the diagonal indicate the square root of the AVE. Pearson's Correlation is significant at the 0.000 level (2-tailed)											

Table 5: Discriminant validity final nine constructs (average variance extracted and correlations)

Analysis and Results

Analytical strategy. To test the developed hypotheses, the analysis used ordinary least squares (OLS) regression-based path analysis to estimate effects of the antecedent, mediators, and moderators (direct and indirect). This type of integration of mediation and moderation analysis, combining mediation and moderation in one model, was developed by Hayes (2018) [20]. The analytical approach for this research used the integration of moderation and mediation analysis by using the Hayes PROCESS Macro within SPSS.

The model defined. This research used the conditional PROCESS model, a combination of moderation with parallel mediators combined into a single integrated model, as reflected in the conceptual form per Figure 2: Model Conceptual Form. For this study the defined model in SPSS is the PROCESS Model Number 76 where the predictor (X) in our model is supply chain disruption orientation (DIS) and the consequent variable (Y) is supply chain resilience (Res). The focal antecedent (X) is supply chain disruption orientation (DIS) for this research. The model included five parallel mediators (Mi): key supplier relationship management (Key), supply chain orientation for benevolence (Ben), risk management infrastructure (RSK), mindful organizing (Min), and supply chain orientation by top management (Mgt) and two moderator variables in the model: supply chain orientation for commitment and norms (Norm) and supply chain orientation for credibility (SCOCR). The model also included all control variables: geography of the supplier, tier-level within the supply chain of the supplier, functional job role of the respondent, and industry of the firm of the respondent (GEO, LVL, RL, and Ind). The model was based on 296 respondents who had completed all the 45 items and had also answered all four control variables (GEO, LVL. RL, and Ind), thus the model dropped 8 respondents that did not include completion of all four of the control variables.

The moderation and mediation model were estimated by PROCESS using Model 76. PROCESS uses ordinary least squares regression for model estimation. The regression coefficients are unstandardized in this model as PROCESS does not generate standardized coefficient where the standardized coefficients are not available for models with moderators using the PROCESS macro in SPSS. The control variables are included as antecedent variables in the PROCESS Model. The natural log of all antecedent variables, mediators, moderators, and the consequent variable were used.

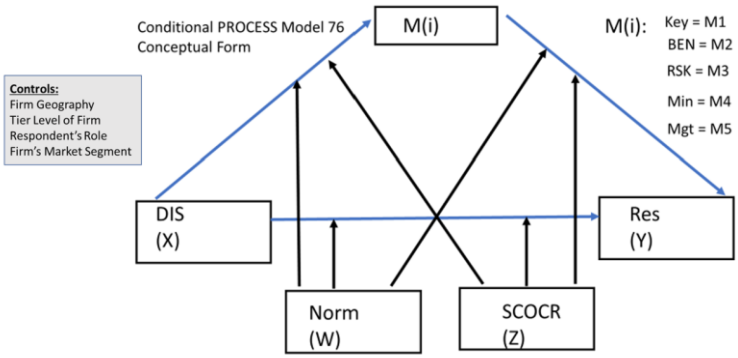


Figure 2: Model Conceptual Form

Model output. The PROCESS model provided an R-Squared value of 0.5312 indicating that 53.12% of the variance of the consequent variable, supply chain resilience (Res), was explained by the variance of the independent variables (p=.0000), per Table 6: Model Summary. One of the predictor variables that had significant influence on supply chain resilience was the moderating variable supply chain orientation for credibility (SCOCR), (beta= 0.5241, p=0.0002), per Table 7: Model Output. Supply chain orientation towards credibility (SCOCR) includes both increased knowledge of products and services and not making false claims on supply chain members. SCOCR has a significant effect on supply chain resilience (p=0.0002). As credibility behavior orientations increase, so does the supply chain resilience. Figure 3: Model Statistical Form illustrates the statistical model and the significant interactions (p<0.0500).

In terms of interactions in the complex model, there were four interactions that had significant effect ($p<0.0500$) on supply chain resilience. Supply chain orientation for norms and commitment (Norm) was included in three of the significant interactions. Supply chain disruption orientation (DIS), the focal antecedent, when interacting with supply chain orientation for norms and commitment (Norm) was significant (beta=0.4432, p=0.0165). In the model, supply chain orientation for norms and commitment also had a significant interaction with mindful organizing (Min) (beta= -0.4241, p=0.0362). Norm also had a significant interaction with supply chain orientation by top management (Mgt) (beta= 0.4021, p=0.0194). There is a second interaction including supply chain orientation by top management (Mgt) which is included in the fourth significant interaction with supply chain orientation for credibility (SCOCR) (beta= -0.4459, p=0.0097). The model output, per Table 7, indicates the significant variables and interactions, Interaction 1, 5, 6, 12 and SCOCR and Lvl.

There is no significance on supply chain resilience consequence when testing the interactions of the focal antecedent (X), supply chain disruption orientation, with the five mediators ($p>0.3000$): key supplier relationship management, supply chain orientation for benevolence, risk management infrastructure, mindful organizing, and supply chain orientation by top management, per the PROCESS tests of X by Mi interactions results per Table 8.

Regarding the effect of the control variables, there was one control variable with a significant effect on supply chain resilience. This was the tier level that the respondent represents and interacts with amongst the multi-layered supply chain delivering their goods and services (beta=-0.1878, p=0.0000). There is a significance on resilience that depends upon the supply chain tier-level at which the respondents' operated at within the relationship to the value chain of the supply chain. The remaining control variables findings indicated that geography (beta=-0.0209, p=0.6201), job

functional role of the respondent (beta=.0458, p=0.1061), and the industry in which the respondent worked (beta= -0.0055, p=0.8788) all had nonsignificant influence on supply chain resilience.

The variance-covariance matrix, per Table 9, shows the variance between pairs of variables in the diagonal (same variable) and the covariance in the off-diagonal pairs. Per Table 9, there is very small change in one variable that is associated with changes in a second variable. The covariance numbers indicated no degree to which two variables are linearly associated. The covariances are not zero, but no greater than 0.0095 for any of the off-diagonal elements that depict the covariances between all variables. The analysis verified that the variables in the model are not strongly correlated (covariance values close to zero), hence are not predictors together in this model. Per Table 9, the variance-covariance matrix provides the statistics in data displaying the variability and covariation.

R	R-sq	MSE	F	df1	df2	p
0.7289	0.5312	0.0922	12.7961	24.0000	271.0000	0.0000

Table 6: Model Summary – Outcome Variable Supply Chain Resilience (Res)

	Full Model		Dyad 1		Dyad 2	
	coeff	**p**	coeff	**p**	coeff	**p**
constant	0.2063	0.0840	**0.3471**	**0.0021**	0.0860	0.7336
DIS	-0.0985	0.5859	-0.1624	0.4026	-0.4595	0.2243
Key	0.3679	0.0992	0.4895	0.0577	0.7267	0.1986
Ben	0.2754	0.0739	0.1979	0.2390	**0.5400**	**0.0393**
RSK	-0.0580	0.6589	-0.1660	0.2450	-0.0450	0.8331
Min	0.1987	0.2682	0.2782	0.1576	0.2664	0.3920
Mgt	0.0396	0.7937	-0.0163	0.9221	-0.5124	0.1280
Norm	-0.1579	0.2611	-0.2546	0.1041	-0.4880	0.1224
Interaction_1 (DIS*Norm)	**0.4432**	**0.0165**	**0.4029**	**0.0357**	**1.0317**	**0.0288**
Interaction_2 (Key*Norm)	-0.2997	0.1798	-0.2746	0.2437	-0.5714	0.3743
Interaction_3 (Ben*Norm)	0.0249	0.8800	0.0054	0.9748	-0.2270	0.5641
Interaction_4 (RSK*Norm)	0.1799	0.2278	0.1960	0.2582	0.1658	0.6269
Interaction_5 (Min*Norm)	**-0.4241**	**0.0362**	-0.3268	0.1515	-0.6131	0.1264
Interaction_6 (Mgt*Norm)	**0.4021**	**0.0194**	**0.4268**	**0.0199**	**1.0377**	**0.0175**
SCOCR	**0.5241**	**0.0002**	**0.5590**	**0.0005**	**0.8994**	**0.0032**
Interaction_7 (DIS*SCOCR)	-0.2628	0.1378	-0.2281	0.2441	-0.5027	0.2228
Interaction_8 (Key*SCOCR)	0.0635	0.7429	-0.1371	0.5291	-0.4172	0.4993
Interaction_9 (Ben*SCOCR)	-0.1520	0.3325	-0.0101	0.9515	-0.1175	0.7938
Interaction_10 (RSK*SCOCR)	0.0251	0.8407	0.1271	0.3912	0.3194	0.2860
Interaction_11(Min*SCOCR)	0.2760	0.1395	0.1297	0.5455	0.3833	0.3426
Interaction_12 (Mgt*SCOCR)	**-0.4459**	**0.0097**	**-0.4172**	**0.0384**	-0.5821	0.1990
LVL	**-0.1878**	**0.0000**	**-0.2543**	**0.0001**	0.1212	0.5553

Table 7: Model Output

	F	df1	df2	p
M1*X	0.9629	1	270	0.3273
M2*X	0.0543	1	270	0.8159
M3*X	0.0823	1	270	0.7745
M4*X	0.1008	1	270	0.7511
M5*X	1.3209	1	270	0.2514

Table 8: Test(s) of X (DIS) by M(i) Interaction

	constant	DIS_ln	Key_ln	Ben_ln	RSK_ln	Min_ln	Mgt_ln	Norm_ln	SCOCR_ln
constant	**0.0142**	-0.0015	-0.0013	-0.0005	-0.0040	-0.0054	-0.0008	-0.0080	-0.0022
DIS_ln	-0.0015	**0.0326**	-0.0048	-0.0025	-0.0002	-0.0038	-0.0036	-0.0008	-0.0020
Key_ln	-0.0013	-0.0048	**0.0494**	-0.0123	-0.0075	-0.0106	-0.0095	0.0020	-0.0009
Ben_ln	-0.0005	-0.0025	-0.0123	**0.0236**	0.0003	-0.0060	-0.0041	-0.0015	0.0013
RSK_ln	-0.0040	-0.0002	-0.0075	0.0003	**0.0172**	-0.0021	-0.0038	0.0016	0.0022
Min_ln	-0.0054	-0.0038	-0.0106	-0.0060	-0.0021	**0.0321**	-0.0038	0.0057	-0.0020
Mgt_ln	-0.0008	-0.0036	-0.0095	-0.0041	-0.0038	-0.0038	**0.0228**	-0.0002	0.0008
Norm_ln	-0.0080	-0.0008	0.0020	-0.0015	0.0016	0.0057	-0.0002	**0.0197**	-0.0091
SCOCR_ln	-0.0022	-0.0020	-0.0009	0.0013	0.0022	-0.0020	0.0008	-0.0091	**0.0193**

Table 9: Covariance matrix of regression parameter estimates

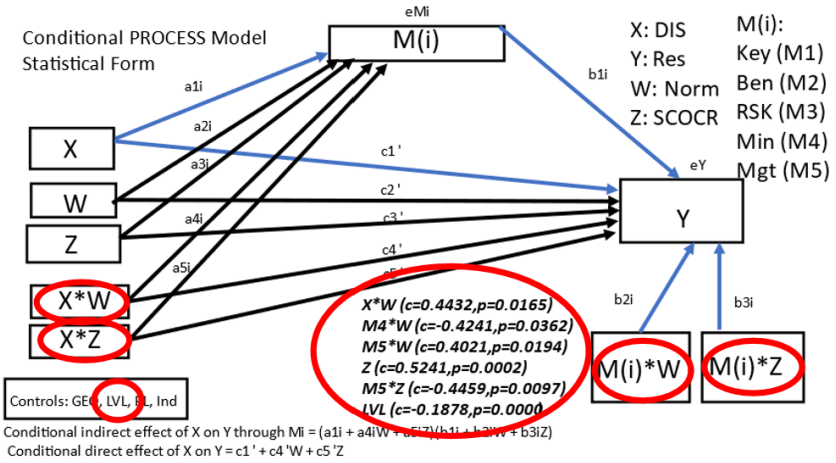


Figure 3: Model Statistical Form

Dyadic model. Using the same PROCESS Model 76, we compared the models between two different dyadic relationships. Model Dyad 1 included the OEM and Tier 1 (External Manufacturer) respondents and then Model Dyad 2 included the Tier 1 (External Manufacturer) and Tier 2 (sub-tier suppliers). The purpose was to compare the dyadic relationships relative to behavior that effects supply chain resilience. The model antecedents, mediators, moderators, and consequent variables were not changed from the original model. But the focus here was on the significance of tier level dyads on supply chain resilience. The N respondents for Model Dyad 1 included 241 and the Model Dyad 2 included 97 respondents.

The PROCESS model provided an R-Squared value of 0.5150 and 0.5723 for Model Dyad 1 per Table 10 and Model Dyad 2 per Table 11 respectively. Supply chain resilience in both dyad relationships, was explained by the variance of the independent variables (p=.0000).

Dyad Description			**Model Summary**						
Dyad	Comparison Pair	Sample Size	R	R-sq	MSE	F	df1	df2	p
Dyad 1	OEM & Tier 1 (EM)	241	0.7176	**0.5150**	0.0843	11.0726	21.0000	219.0000	0.0000

Table 10: Model summary – Dyad 1 (OEM and Tier-1/EM)

Dyad Description			**Model Summary**						
Dyad	Comparison Pair	Sample Size	R	R-sq	MSE	F	df1	df2	p
Dyad 2	Tier 1 (EM) & Tier 2 (subtier)	97	0.7565	**0.5723**	0.0952	4.7783	21.0000	75.0000	0.0000

Table 11: Model summary – Dyad 2 (Tier-1/EM and Tier-2/Sub-tier Assembly)

Both models explain greater than 50 % of the variance of the respective models. Both Dyad 1 and Dyad 2 Models include significant variables and interactions, per Table 7: Model Output.

Benevolence is a significant contribution to supply chain resilience in Dyad 2 relationships (b=0.5400, p= 0.0393) singled out among the five mediators, per Table 7. Supply chain orientation for credibility is significant contribution to supply chain resilience in Dyad 2 relationship (b=0.8994, p=0.0032), one of the two moderators. The same is true for Dyad 1 (p=0.0005) for supply chain orientation for credibility significant effect on supply chain resilience, per Table 7. The two interactions that are significant contributions to supply chain resilience in Dyad 2 depend upon the supply chain orientation for norms and commitments (Norm), products Dis*Norm and Mgt*Norm (p=0.0288 and p=0.0175 respectively). Both interactions are also significant in Dyad 1 (p=0.0357 and p=0.0199 respectively). Regardless of dyad there is significance in these very same interactions.

However, supply chain benevolence mediation is not significant within the dyadic relationship between the OEM and the tier 1 external manufacturing supplier, Dyad 1. The OEM and EM relationship, Dyad 1, is not significant for any mediator or moderator variables, but is significant at the same interactions as Dyad 2, except there is one added significant interaction that is supply chain orientation by top management (Mgt) with supply chain orientation for credibility (SCOCR) represented as the product of variables, Mgt*SCOCR (p=0.0384). This interaction includes a mediation variable and a moderation variable interaction in the product. The interaction of the supply chain orientation of top management and supply chain orientation for credibility is significant within the relationship between the OEM and external manufacturers (tier 1), Dyad 1, refer to Table 7.

The further away from the OEM in the supply chain, is the greater the significance on supply chain orientation for benevolence (Ben) as it is a significant effect on supply chain resilience (Res) in the Dyad 2 model (p=0.0393). This mediator, Ben, is not significant in the Dyad 1 Model. Between the OEM and EM suppliers (Tier 1) the control variable LVL, supplier tier-level within the supply chain, is significant (p=0.0001), but is not in Dyad 2 Model. The coefficient of the LVL is negative, indicating that this dyad relationship decreases resilience. The supply chain level is not a significant contribution to the model in Dyad 2 as the further away from the OEM in the supply chain, the lesser the influence of different level/tiers is i.e., not significant (p=0.5553), refer to Table 7. The full model results compared to Dyad 1 and Dyad 2 are summarized per Table 7: Model Output.

Results and theoretical implications. This research leveraged existing research relative to the supply chain disruption orientation (SDRO) construct that has been associated with an increase in supply chain resilience. This research included the addition of antecedents contributing to the level of supply chain resilience beyond SDRO. The leveraged prior research included supply chain orientation (SCO), key supplier relationship management (KSRM), risk management infrastructure (RMI), combined with mindful organizing (MO) as factors affecting the level of supply chain resilience (SCR) during disruptive events. These five independent variables are existing constructs from prior research with associated verified scales, but including all of these in one model was novel including the combination of all five as mediators or moderators within the SDRO and SCR relationship. All these complex supplier relationship factors were included in a survey of 52 items to learn more about supply chain relationships and interconnectedness relative to supply chain resilience.

The final analysis increased the constructs from six to nine and reduced total items to 45, based on loading. The consequent variable (SCR) held together (4 items) renamed Res in the final model, mindful organizing held together (MO) renamed Min, key supplier relationships (KSRM) held together renamed Key, and risk management infrastructure (RMI) held together renamed RSK. There were constructs that did not hold together, with recoding and some items dropped. The focal antecedent, supply chain disruption orientation (SDRO) did not hold together as it dropped items and added another from the SCO construct renamed DIS in the final model. The SCO construct did not stay as one construct, but broke apart and loaded as four separate independent variables forming new constructs in the final model based upon loading as shown in Table 12: Measurement Scales.

The newly formed constructs replaced SCO aggregate based on the factor loading into newly named **SCOCR, Ben, Norm, and Mgt**. SCOCR is credibility orientation assuring firms in the supply chain are not making false claims amongst supply chain members and depends upon each firm to be knowledgeable about their own products. The Ben variable is when a supply chain firm has concern about the welfare of the supply chain including depending upon supply chain members' support and responding with understanding when there are problems. Ben variable requires that decisions and actions count on supply chain members to consider how other supply chain members are affected. The Norm variable is defending the members of the supply chain when they are trusted, being patient when mistakes are made qualified to not be "repeated" mistakes, but the essence of Norm is a cooperative culture. The Mgt variable includes top management repeatedly telling employees that their survival and success depends on adapting, building relationships, sharing information, risk and sharing reward along with education about supply chain management.

***Independent variables, interactions, moderation implications*.** Supply chain orientation for credibility (SCOCR) is a significant (beta=0.5241, p=0.0002) effect on resilience. Also, when firms are not making false claims and knowledgeable about their products (SCOCR) interacts with supply orientation for top management (Mgt*SCOCR) there is significant negative effect on resilience (beta=-0.4459, p=0.0097). When top managers repeatedly remind employees about supply chain criticality (Mgt) and there is a culture present that is making sure that its supply chain professionals are not making false claims (SCOCR), there is conflicting behavior in the supply chain members creating a significant interaction that negatively affects resilience.

There are three interactions with supply chain orientation for norms and commitment (Norm) that have a significant effect on supply chain resilience. Firms with Norm defend their supply chain members, are patient when supply chain members make mistakes (not repeatedly), and are willing to make cooperative changes with supply chain members. Two of the three significant Norm interaction behaviors should be built into a firm's supply chain function and culture (DIS and Mgt), when there is the presence of Norm, to build resilience. One interaction behavior (Min) should be avoided (in the presence of Norm) as it takes away from resilience as an interaction: DIS*Norm (beta=0.4432, p=0.0165), Min*Norm (beta=-0.4241, p=0.0362), Mgt*Norm (beta=0.4021, p=0.0194).

When "Norm" behaviors interact with supply chain disruption orientation (DIS*Norm), there is a significant positive effect on supply chain resilience (beta=0.4432, p=0.0165). Firms with DIS are always alert for disruptions, recognize that disruptions are always looming, and use disruptions to improve are the firms that have better supply chain resilience as confirmed in this research as an interaction and by Ambulkar et al. (2015) [9] as an independent construct. When firms have a heightened awareness of disruptions (DIS) and interact with a firm's belief that their supply chain

members must work together to be successful (Norm), there is increased supply chain resilience. Therefore, firms should drive both orientations to increase stronger supply chain resilience (Dis*Norm). A key supply chain firm message is that recognizing disruptions are forthcoming is not enough by itself, but the firm must in addition, be honest in all dealings in the supply chain. This combined interaction is cultural and learned behaviors for supply chain firm investments when hiring and retaining supply chain professionals.

The second significant interaction with Norm is when there is the presence of mindful organizing (Min) in a firm (Min*Norm), there is a significant negative effect on supply chain resilience (beta = -0.4241, p=0.0362). Mindful organizing when combined at the same time with Norm in a firm is pulling time and energy internal to the firm and not looking outward to the supply chain nodes. It may also lead to "analysis paralysis" taking away from supply chain resilience.

The third significant interaction with supply chain orientation with Norm is when there is supply chain orientation for top management (Mgt) in a firm (Mgt*Norm), there is a significant positive effect on supply chain resilience (beta=0.4021, p=0.0194). The interaction with top managers repeatedly 'telling' their employees about supply chain criticality has an additive effect on resilience when included with firms that have the presence of Norm at the same time. The 'telling' by top management, combined with a firm that is committed to its suppliers, is additive to supply chain resilience. The key message to management is that top management emphasizing (Mgt) survival requirements (adapting, building, sharing with supply chain) must be complemented with a culture of patience, defending, and willing to make cooperative changes with supply chain members (Norm) to strengthen supply chain resilience. This is a culture combination that is required with top management enabling norm concurrently.

Managerial implications. To create more resilience in supply chains, managers should continue to develop a sense for disruption (orientation) (DIS) within their culture that requires employees specifically to be alert and on the lookout for the next disruption. At the same time, firms require a culture defending supply chain members when outsiders criticize them (if trusted) and be patient when these same suppliers make mistakes (Norm). There is an increase in resilience due to these interactions (both must be present in the firm). Do not invest in Mindful organizing behaviors in the presence of Norm culture as this pulls the firm's time and energy to internal analysis which takes away from resilience. Managers need to keep their supply chain professionals working with their supply chain, not analyzing internally as a distraction away from the supply chain nodes, when there is Norm present. An additive action by top management when there is Norm present should include repeatedly 'telling' their employees about supply chain criticality (Mgt) as it has an additive effect to resilience, when there is the presence of Norm.

Firms investing in resilience programs should continue to work with their supply chain professional's orientation for credibility (SCOCR). Supply chain orientation for credibility influences resilience on its own without interacting with other relationship behaviors. Credibility is a critical anchor to supply chain resilience regardless of tier level and dyad relationship as it is a significant and positive beta for all models.

The firm's position in the supply chain relative to tier-level in the supply chain influences supply chain resilience. Firms that are not the OEM in the supply chain should be focusing on their investment in benevolence with their supply partners they interact with on regular basis, either a tier up or down in their supply chain interactions and relationships. The further away suppliers are from the OEM and EM, the more their resilience depends upon the supply chain orientation for benevolence (Dyad 2) as a significant impact on resilience (beta=0.5400, p=0.0393). Regardless

of their level, the investment in Norm alongside two of the focal antecedents, DIS and Mgt, there is positive effect on resilience. Therefore, supply chain tiers must have both to influence positively to supply chain resilience regardless of the respective tier in the supply chain.

Limitations and suggestions for further research. This research was conducted via on-line survey primarily distributed to a high-technology industry survey population, which could bias the results towards this market segment. The next step could be to expand this to other industries for comparison. Also, the quantity of contacts at tiers below the OEM was limited in distribution, which affects the Dyad 2 results. Another survey could be completed to target sub-tier suppliers and not OEMs. More research is required at the next dyad tier-levels beyond OEMs, tier-1 and tier-2 as their relationships are affected by disruptions and the entire supply chain relies upon resilience at all tier-levels for optimized delivery of goods and services meeting customer expectations.

CONCLUSION

The complexity of this model mimics supply chain relationships in complexity. The recent pandemic proved there were relationships that created resilience and others that did not across many industries over the past years. This research indicates there is not one relationship behavior that stands on its own but multiple behaviors that interact in an additive novel way to support supply chain resilience or the collective presence may take away from resilience. Firms may have one or two behaviors, but not all that were measured in this research, creating sub-optimal resilience factors during disruptions effecting the supply chain. Finally, firms need to be aware that benevolence is a differentiator within the tier 1 and tier 2 relationships affecting resilience enabling these two tiers to manage and navigate around disruptions. Tier-1 and tier-2 suppliers rely upon genuine concern for one another's welfare when making important decisions, response

to one another with understanding when dealing with supply chain problems amongst one another, and these supply chain members depend on supply chain members' support when it comes to things that are important to either tier. Benevolence in this model looks to the future and expects that a dyad can count on their supply chain members to consider how their decisions and actions will affect one another. Tier to tier interconnectedness depends upon each firm's benevolence to increase supply chain resilience. This disruption and resilience research provides new supply chain insight through the lens of relationships and interconnected behaviors.

All scales were 7-point Likert with anchors "Strongly disagree" and "Strongly agree"

Construct	Indicator Element	Standard Deviation	Loading	Average Variance (AVE)	Cronbach's alpha (construct reliability)	Composite Reliabilities (CR)
Supply Chain Resilience (SCR) (Adapted from Ambulkar, Blackhurst, & Grawe, 2015).	**SCRALL (4x items)- Res (Y)**			**0.598**	**0.880**	**0.856**
We are able to cope with changes brought by the supply chain disruption	SCR_1	1.279	0.771			
We are able to adapt to the supply chain disruption easily	SCR_2	1.483	0.805			
We are able to provide a quick response to the supply chain disruption	SCR_3	1.482	0.787			
We are able to maintain high situational awareness at all times	SCR_4	1.430	0.727			
Supply chain disruption orientation (SDRO) (Adapted from Bode et al. (2011) and Ambulkar et al. (2015))	**DISRUPT (4x items)- DIS (X)**			**0.431**	**0.694**	**0.752**
We feel the need to be alert for possible supply chain disruptions at all times	SDR_O1	0.959	0.650			
Supply chain disruptions show us where we can improve	SDR_O2	0.961	0.671			
We recognize that supply chain disruptions are always looming	SDR_O3	1.106	0.675			
We believe our supply chain members must work together to be successful	SCO_COMM_and_NORM_4	0.984	0.629			
We think a lot about how a supply chain disruption could have been avoided **(D)**	*SDR_O4*		*0.443*			
After a supply chain disruption has occurred, it is analyzed thoroughly **(D)**	*SDR_O5*		*0.463*			
Construct	**Indicator Element**	**Standard Deviation**	**Loading**	**Average Variance (AVE)**	**Cronbach's alpha (construct reliability)**	**Composite Reliabilities (CR)**
Supply chain orientation (SCO) (Adapted from Min & Mentzer, 2004)						
Supply chain orientation (SCO) - Credibility (Adapted from Min & Mentzer, 2004)	**SCOCRED (2x items)- SCOCR (Z)**			**0.559**	**0.710**	**0.716**
Promises made to our supply chain members by our business unit are reliable **(D)**	SCO_CRED_1		*0.469*			

Our business unit is knowledgeable regarding our products and/or services when we are doing business with our supply chain members	SCO_CRED_2	1.062	0.707			
Our business unit does not make false claims to our supply chain members	SCO_CRED_3	1.213	0.786			
Our business unit is not open in dealing with our supply chain members **(D) (R)**	SCO_CRED_4 (CRED_4_Recoded)		0.675			
Supply chain orientation (SCO) -Benevolence (Adapted from Min & Mentzer, 2004)	**SCOBEN (4x items)- Ben (M2)**			**0.513**	**0.874**	**0.808**
When making important decisions, our supply chain members are concerned about our welfare	SCO_BEN_1	1.396	0.703			
When we share our problems with our supply chain members, we know they will respond with understanding	SCO_BEN_2	1.200	0.758			
In the future we can count on our supply chain members to consider how their decisions and actions will affect us	SCO_BEN_3	1.337	0.712			
When it comes to things that are important to us, we can depend on our supply chain members' support	SCO_BEN_4	1.090	0.690			
Supply chain orientation (SCO) - Commitment (Adapted from Min & Mentzer, 2004)	**SCONorm (3x items)-Norm (W)**			**0.466**	**0.773**	**0.719**
We defend our supply chain members when outsiders criticize them, if we trust them	SCO_COMM_and_NORM_1	1.212	0.557			
We are patient with our supply chain members when they make mistakes that cause us trouble but are not repeated	SCO_COMM_and_NORM_2	1.232	0.802			
Supply chain orientation (SCO) - Norms (Adapted from Min & Mentzer, 2004)						
Our business unit is willing to make cooperative changes with our supply chain members	SCO_COMM_and_NORM_3	1.150	0.667			
We view our supply chain as a value-added piece of our business **(D)**	*SCO_COMM_and_NORM_5*		*0.365*			
Supply chain orientation (SCO) - Compatibility (Adapted from Min & Mentzer, 2004)						
Our business units' goals and objectives are consistent with those of our supply chain members **(D)**	*SCO_COMM_and_NORM_6*		*0.436*			
Our CEO and CEOs of our supply chain members have similar operating philosophies **(D)**	*SCO_COMM_and_NORM_7*		*0.409*			
Supply chain orientation (SCO) -Top Management Support (Adapted from Min & Mentzer, 2004)	**SCOMgt (5x items)- Mgt (M5)**			**0.508**	**0.901**	**0.836**
Top managers repeatedly tell employees that this business unit's survival depends on its adapting to supply chain management	SCO_Top_MGT_1	1.54	0.738			
Top managers repeatedly tell employees that building, maintaining, and enhancing long-term relationships with our supply chain members are critical to this business unit's success.	SCO_Top_MGT_2	1.51	0.703			
Construct	**Indicator Element**	**Standard Deviation**	**Loading**	**Average Variance (AVE)**	**Cronbach's alpha (construct reliability)**	**Composite Reliabilities (CR)**
Top managers repeatedly tell employees that sharing valuable strategic/tactical information with our supply chain members is critical to this business unit's success	SCO_Top_MGT_3	1.509	0.798			
Top managers repeatedly tell employees that sharing risk and rewards is critical to this business unit's success	SCO_Top_MGT_4	1.532	0.742			
Top management offers various education opportunities about supply chain management	SCO_Top_MGT_5	1.675	0.559			

Mindfulness (formerly Safety) Organizing Scale (MO) (Adapted from Vogus & Sutcliffe, 2007, 2012).	**Mind (8x Items)- Min (M4)**			**0.485**	**0.915**	**0.882**
We have a good "map" of each person's talents and skills	MO_1	1.392	0.694			
We talk about mistakes and ways to learn from them	MO_2	1.352	0.679			
We discuss our unique skills with each other so that we know who has relevant specialized skills and knowledge	MO_3	1.479	0.746			
We discuss alternatives as to how to go about our normal work activities	MO_4	1.304	0.757			
When discussing emerging problems with co-workers, we usually discuss what to look out for	MO_5	1.181	0.683			
When attempting to resolve a problem, we take advantage of the unique skills of our colleagues	MO_6	1.226	0.735			
When errors happen, we discuss how we could have prevented them	MO_7	1.265	0.704			
When a crisis occurs, we rapidly pool our collective expertise to attempt to resolve it	MO_8	1.088	0.551			
Key Supplier Relationship Management Scale - (KSRM) (Adapted from Crnjak-Karanovic, 2012).	**KeySup (12x items)- Key (M1)**			**0.423**	**0.935**	**0.898**
Key Supplier Relationship Management Scale - (KSRM) -Planning (Adapted from Crnjak-Karanovic, 2012).						
We specify our relationship with every key supplier.	KSRM_1	1.238	0.694			
We analyze with which suppliers we will do business in long term.	KSRM_2	1.210	0.675			
We define our relationship with every key supplier according to the strategic goals of the company.	KSRM_3	1.163	0.728			
We use segmentation principles to divide strategic and transactional suppliers.	KSRM_4	1.191	0.636			
Key Supplier Relationship Management Scale - (KSRM) Implementation - (Adapted from Crnjak-Karanovic, 2012).						
We strive to consolidate our supplier base	KSRM_5	1.351	0.650			
We pay attention to the activities and initiatives that will enhance the performance of our key suppliers.	KSRM_6	1.245	0.709			
We collaborate with our key suppliers in the development of new products that fit our business needs.	KSRM_7	1.145	0.618			
We consider our key suppliers as an extension to our manufacturing processes.	KSRM_8	1.193	0.587			
We use information technology to achieve integration with our key suppliers.	KSRM_9	1.232	0.635			
Key Supplier Relationship Management Scale - (KSRM) -Control - (Adapted from Crnjak-Karanovic, 2012).						
We employ control mechanisms to assess the performance of our key supplier relationships	KSRM_10	1.166	0.638			
The control of key supplier relationships is done according to the strategic goals of our company	KSRM_11	1.134	0.668			
We undertake corrective measures if we determine deviations in our relationship with key suppliers	KSRM_12	1.099	0.547			
Risk Management Infrastructure Scale – (RMI) (Adapted from Ambulkar et al., 2015)						
Construct	**Indicator Element**	**Standard Deviation**	**Loading**	**Average Variance (AVE)**	**Cronbach's alpha (construct reliability)**	**Composite Reliabilities (CR)**
Risk management infrastructure- (RMI)	**RskInfr (3x items)- RSK (M3)**			**0.518**	**0.836**	**0.763**
We have a department to manage supply chain risks and disruptions	RMI_1	1.644	0.688			
We have KPI and metrics to monitor supply chain risk	RMI_2	1.243	0.707			

We have information systems in place to manage supply chain risks and disruptions	RMI_3	1.471	0.762			

R- reverse coded-item,

D-Dropped in purification Process

AVE: Average Variance Expected >.50 or CR >.70

Table 12: Measurement Scales

REFERENCE LIST

[1] Sheffi, Yossi & Rice, James B. Jr., A Supply Chain View of the Resilient Enterprise. MIT Sloan Management Review. Vol. 47 No. 1, 41-48, Fall 2005.

[2] Blackhurst, J., Dunn, K., & Craighead, C.W., 2011. An empirical derived framework of global supply resiliency. Journal Business Logistics, 32 (4), 374-391.

[3] Manuj, Ila & Mentzer, John T., (2008), "Global Supply Chain Risk Management", Journal of Business Logistics, Vol. 29, No.1, 2008.

[4] Sheffi, Yossi (2007). *The Resilient Enterprise: Overcoming Vulnerability for Competitive Advantage.* The MIT Press. Cambridge, MA.

[5] Christopher, Martin & Peck, Helen. Building the Resilient Supply Chain. International Journal of Logistics Management, Vol. 15, No., 2, pp 1-13, 2004.

[6] De Loach, J. W. (2000). Enterprise-wide Risk Management: Strategies for linking risk and opportunity. London: Financial Times/Prentice Hall.

[7] Craighead, C. W., Blackhurst, J., Rungtusanatham, M. J., & Handfield, R. B. 2007. The severity of supply chain disruptions: Design characteristics and mitigation capabilities. Decision Science, 38: 131-156.

[8] Bode, C., Wagner, SM., Petersen, K.J., & Ellram, LM., 2011. Understanding responses to supply chain disruptions: insights from information processing and resource dependence perspectives, The Academy of Management Journal, August 2011, Vol. 54, No. 4, pp. 833-856.

[9] Ambulkar, Saurabh, Blackhurst, Jennifer, & Grawe, Scott. Firm's Resilience to Supply Chain Disruptions: Scale Development and Empirical Examination. Journal of Operations Management, 33-34 (2015) 111-122.

[10] Holling, C. S. (1973). Resilience and Stability of Ecological Systems. Annual Review of Ecology and Systematics, 4, 1–23.

[11] Daft, Richard & Weick, Karl E., Toward a Model of Organizations as Interpretation Systems. Academy of Management Review, Vol.9, No.2. April 1, 1984.

[12] Miocevic, Dario, & Crnjak-Karanovic, Biljana, The mediating role of key supplier relationship management practices on supply chain orientation – the organizational buying effectiveness link. Industrial Marketing Management. Industrial Marketing Management 41 (2012) 115-124.

[13] Mentzer., J.T. DeWitt, W., Keebler, J.S., Min, S., Nix, N.W., Smith, C.D., & Zacharia, Z.G. Defining Supply Chain Management, Journal of Business Logistics, Vol 22, No. 2, 2001, pp 1-26.

[14] Min. S., & Mentzer, J.T. (2004). Developing and measuring supply chain management concepts, Journal of Business Logistics 25 (1), 63-99.

[15] Vogus, Timothy J., & Sutcliffe, Kathleen M., The Safety Organizing Scale: development and validation of a behavioral measure of safety culture in hospitals nursing units. Medical Care, Volume 45, Number 1, pgs 46-54, January 2007.
[16] Vogus, Timothy J., and Sutcliffe, Kathleen M., Organizational Mindfulness and Mindful Organizing: A reconciliation and path forward. Academy of Management Learning & Education. Vol. 11, No.4, pp 722-735. December 2012.
[17] Berente, N., Lyytinen, K., Yoo, Y., & King, J. L. (2016). Routines as shock absorbers during organizational transformation: Integration, control, and NASA's enterprise information system. Organization Science, 27(3), 551-572.
[18] Esper, Terry L., Deffee, Clifford C., Mentzer, J.T., "A framework of supply chain orientation", The International Journal of Logistics Management, Vol 21, No. 2, 2010. pp161-179.
[19] Weick, K. E., & Sutcliffe, K. M. (2006). Mindfulness and the Quality of Organizational Attention. *Organization Science*, *17*(4), 514–524.
[20] Hayes, Andrew F. (2018), Introduction to Mediation, Moderation, and Conditional Process Analysis (2nd edition). The Guilford Press.

Effects of Consumer Showrooming and Supplier Encroachment on an Omnichannel Retailer

(Authors' names blinded for peer review)

The recent burgeoning popularity of retail e-commerce has enabled many manufacturers to reach diverse consumer segments through online direct channels. This practice, known as supplier encroachment, allows an upstream manufacturer to compete with his downstream retailer in the retail market. In recent years, most retailers have adopted an omnichannel strategy, with a brick-and-mortar (B&M) channel and an online channel of their own. Consumers may prefer one online channel over another, but on either online channel, they are unable to ascertain the product fit with respect to non-digital attributes, unless they utilize the B&M store for evaluation. This often results in a practice called consumer showrooming, where some consumers visit the B&M store to evaluate a product, but then buy it from a different online store. In this paper, we study how the interaction of these two effects impacts the omnichannel retailer and her manufacturer.

Note: Bulk of the work in this paper has been done by a student.

Keywords: Supplier Encroachment, Showrooming, Omnichannel Retail, Analytical Modeling

1. Introduction

With the growing prominence of e-commerce as a retail channel in recent years, companies have increasingly adopted online selling strategies (Caro et al. 2020). In the United States of America alone, consumers spent USD 601.7 billion in e-commerce sales in 2019, an increase of 14.9% from USD 523.64 billion in 2018 (Young 2020), while total retail spending went up 2.8%, from USD 5.32 trillion in 2018 to USD 5.47 trillion in 2019 (O'Connell 2020). E-commerce sales in 2019 accounted for 11% of total retail sales, up from less than 10% a year ago (Keyes 2020).

Amid this widespread popularity of e-commerce, online retailers, such as Amazon.com, have grown extraordinarily, while traditional brick and mortar (B&M) retailers have expanded their omnichannel presence (Bell et al. 2018). At the same time, consumer brand manufacturers have grown at a much faster rate online than retail chains and catalog stores (Davis 2019). The e-commerce boom has widely enabled suppliers, such as Apple, Google, Nike, and Ralph Lauren, to *encroach* the retail space, and directly sell to diverse end-consumer segments (Guan et al. 2019, Xia and Niu 2019). The practice of *supplier encroachment*, where an upstream supplier competes with its retailer in the retail market, has been of interest in the supply chain literature in recent years (Chiang et al. 2003, Cattani et al. 2006, Arya et al. 2007, Li et al. 2014, 2015, Huang et al. 2018). However, while most of the attention in this body of literature has been on the effect of increasing competition at the retail level, and on the increasing supply chain coordination achieved by the supplier, an important aspect of consumer behavior, *showrooming*, has not been explored in detail.

Distinct from an online-first retailer creating physical showrooms for the purpose of displaying products (Arya and Mittendorf 2018, Bell et al. 2018), *consumer showrooming* is a strategic behavior demonstrated by certain consumer segments. While e-commerce provides many benefits to consumers, they are often unable to experience the tangible properties of a product when they exclusively evaluate and buy it online. When the product has a significant amount of non-digital attributes, as defined by Lal and Sarvary (1999), this results in product fit uncertainty when consumers transact exclusively online (Hong and Pavlou 2014). These attributes cannot be completely assessed by the consumer unless the product is examined in person, even after the consumer has exhausted online options, such as product reviews (Chiang et al. 2003, Balakrishnan et al. 2014). In order to avoid the loss of utility resulting from a misfit product, some consumers engage in the practice of *consumer showrooming*, where they first visit a physical (i.e., B&M) retail store to find the product(s) that fit best, but then buy the product(s) from a different, online, retail store. This phenomenon, sometimes known as *browse-and-switch*, has been well-documented in the popular press (Zimmerman 2012, Pearson 2017, Cain 2018), and it has piqued interest in the academic literature in recent years (Balakrishnan et al. 2014, Mehra et al. 2018).

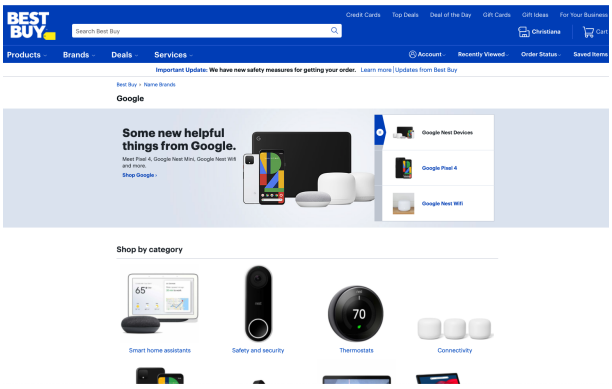
As suppliers become more inclined to take advantage of the advent of e-commerce by opening direct online channels, their omnichannel retail partners face a two-pronged challenge. First, the retailer faces increased competition at the retail echelon due to the supplier's encroachment. Secondly, the retailer's B&M store faces the prospect of becoming a showroom for certain strategic consumer segments. While the literature has considered each of these effects individually, the simultaneous interaction of these two phenomena has not been extensively studied. In this paper, we investigate a scenario where an upstream supplier, by establishing an online channel, can enter the retail market and encroach upon its downstream retailer, while encouraging certain consumers to "showroom", or "browse-and-switch".

1.1. Motivation

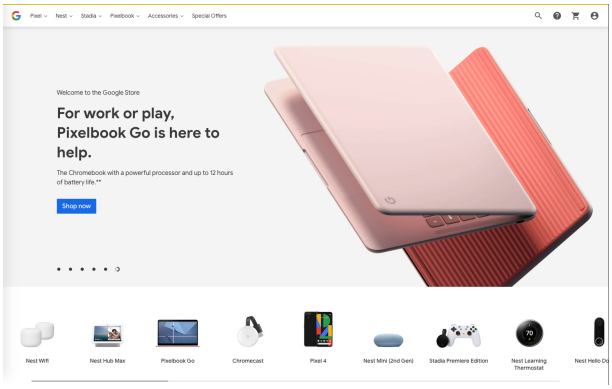
Let us now motivate the discussion with the following example. While Google is known as a leading software provider, it has gained prominence in recent years as a maker of a diverse portfolio of hardware (Hollister 2020). Google sells its hardware products, such as the Google Nest and Google Pixel devices, through omnichannel retailers like Best Buy (with both B&M and online stores, as well as its exclusively-online shop, the Google Store (see Figure 1).

In this situation, consumers have three channels – two online, and one physical – through which they could buy a Google product. Therefore, Google, the manufacturer, has encroached upon the retail space formerly occupied by Best Buy, the retailer. Further, while consumers can evaluate the product online to a certain degree by reading reviews and technical specifications, they cannot

(a) Best Buy B&M Store, Moon (2016)



(b) Best Buy Online Store, Google Brand Store (2020)



(c) Google Online Store, Google Store (2020)

Figure 1 Google Products Sold Through Omnichannel Retailer (Best Buy) Stores and Own Online Store

personally experience and ascertain the fit of non-digital attributes, which is critical for such sophisticated electronic equipment (Balakrishnan et al. 2014, Mehra et al. 2018). Therefore, some consumers visit the B&M store of Best Buy to evaluate the Google device in person. However, not all consumers may buy the Google device at the store. Rather, they are likely to search for better prices online, even while inside the B&M store (Cao 2019). Some consumers may end up purchasing from Google's online store if they find a better price there, resulting in consumer showrooming.

The combined effect of showrooming and encroachment may result in Best Buy losing consumers to Google. However, Google benefits from Best Buy selling Google devices, as consumers may sometimes be more inclined to buy from Best Buy's online store. Consumers may have reasons other than price for their channel preferences (Dumrongsiri et al. 2008, Li et al. 2015). Further, Google may not be as efficient in selling products as Best Buy (Arya et al. 2007, Hotkar and Gilbert 2021). Finally, Best Buy provides the much-needed showroom for Google, which some other manufacturers need to invest in, when they do not have a retail channel partner Bell et al. (2018).

The interaction of these effects results in interesting dynamics that arise in practical situations as described above. However, this strategic interaction (e.g., between Google and Best Buy)

has not been studied extensively in marketing, retail, or supply chain management literature. Our paper bridges this existing gap in the literature, while providing managerial insights for practitioners. In this paper, we study how changing different parameters (such as the operating costs of the manufacturer or the retailer) impacts the online manufacturer and the omnichannel retailer. In the following section, we discuss the streams of literature from where we take our inspiration and how we contributed to these streams of literature.

2. Literature Review

Our research is related to two streams of literature: (1) consumer showrooming and (2) supplier encroachment. In this section, we mention the relevant works from each stream and discuss our contributions to both streams.

2.1. Consumer Showrooming

With the increasing popularity of e-commerce, the practice of consumer showrooming has been increasing and has steadily gained attention from practitioners (Zimmerman 2012, Pearson 2017, Cain 2018). The concept of free-riding, which happens when consumers use one retailer for presales activities, such as getting a salesperson to help with trying different products, and potentially a different retailer for the actual purchase, has been explored in the marketing literature (Singley and Williams 1995, Carlton and Chevalier 2001, Antia et al. 2004). While most of the research has focused on the negative effects of free-riding, it has also been shown to benefit the retailer who invests in presales activities (Wu et al. 2004, Shin 2007). However, this body of research considers retailers of a similar nature, e.g. both retailers operating B&M stores, such that the cost of evaluating a product at either retailer is the same for a consumer. This research also assumes that the competing retailers have no other relationship, whereas we consider a vertical supply chain relationship between the two retailers.

More recently, Balakrishnan et al. (2014) establish that showrooming (or browse-and-switch) behavior is exhibited by consumers when they face uncertainty about product valuation and different costs for utilizing the online and B&M channels, and this behavior intensifies the competition between channels.

Many studies in this line of work explore the arrangement where there is a BM retailer and an online retailer and study how consumer showrooming affects the retail BM stores. Balakrishnan et al. (2014) show that showrooming intensifies competition between physical and online retailers. Mehra et al. (2018) find that the BM store's profit lowers due to showrooming. They discuss different ways the BM store may respond to this phenomenon and try to recover their loss. For example, they may include store brands or exclusivity product brands. Another study by Jing (2018) discusses different strategies adopted by online and physical retailers that aggravate or

alleviate consumer showrooming. By examining how consumers' showrooming behavior interacts with their *webrooming* behavior, this study shows that when webrooming resolves partial consumers fit uncertainty, it may result in an increase in profit for both the physical retailer and the online retailer. Our study contributes to the aforementioned literature by considering the online retailer as a manufacturer and an upstream supplier to the physical retailer, and by including the concept of consumer showrooming behavior in this setting.

2.2. Supplier Encroachment

A number of papers in this stream study how encroachment lowers wholesale prices and benefits downstream resellers. Cattani et al. (2006) analyze pricing strategy in a situation where a manufacturer, having a traditional channel retailer, opens an online direct selling channel. They find that the downstream retailer and consumers always prefer the manufacturer to select the equal-pricing strategy that optimizes profits for the manufacturer over equal-pricing strategies where the manufacturer keeps the wholesale or retail prices unchanged. Arya et al. (2007) study supplier encroachment and show that retailers benefit from this situation in that the supplier often reduces the wholesale price to ensure demand for their product. The lower wholesale price and increased downstream competition mitigate double marginalization and increase total supply chain profit. Chiang et al. (2003) analyze manufacturer encroachment considering consumers' preference for the manufacturer's direct channel. They investigate a price-setting game between a manufacturer and its independent retailer and find that the encroachment benefits the retailer when the consumers' preference for the manufacturer's channel is medium. Zhang et al. (2020) study manufacturer encroachment with informative advertising and show that both manufacturer and retailer benefit if the advertising is controlled by either the manufacturer or by both. Our study contributes to this section of literature by considering that i) the manufacturer does not have a physical store and only operates through an online store and ii) consumers engage in showrooming before purchasing. To the best of our knowledge, our work is the first study to investigate the combined effect of consumer showrooming and supplier encroachment. In the succeeding sections, we first describe different aspects of our model, then we discuss our results, and finally, we conclude with managerial insights derived from our analysis.

3. The Model

In this section, we discuss different aspects of our model, including the supply chain structure, the utility derived from the product, the heterogeneity among consumers, and the consumer decision process. We first discuss the supply chain structure.

3.1. The Supply Chain

As discussed in Section 1, the increasing popularity of e-commerce has resulted in many consumer brands selling their products through a traditional retail channel as well as their own online stores, while retailers have increasingly adopted an omnichannel strategy (Bell et al. 2018, Davis 2019, Guan et al. 2019). Therefore, we consider a supply chain consisting of a manufacturer (he), who sells the product described above through (a) an omnichannel retailer (she), who has a physical (B&M) store and an online store, and (b) the manufacturer's own online store. We denote the retailer's B&M store with subscript S, her online store with SO, and the manufacturer's online store with O.

One of the most commonly used contracts between manufacturers and retailers is the wholesale price contract (see e.g., Li et al. 2014, Roy et al. 2019), the simplicity of which makes it an increasingly popular choice for retailers (Cui et al. 2020). Consequently, while other forms of contracts also exist in the industry (see, e.g., Cachon (2003), we specifically focus on wholesale price contracts between the manufacturer and the omnichannel Retailer. In such contractual arrangements, the manufacturer offers a wholesale price w, to which the retailer responds with her purchase quantity q_R (Arya et al. 2007, Li et al. 2014, Ha et al. 2016), which is the sum of the quantities q_S and q_{SO}, intended for sale through her B&M and online stores, respectively. The manufacturer then determines the selling quantity q_O and price p_O in his own channel (Arya et al. 2007, Li et al. 2015, Hotkar and Gilbert 2021).

Following common practice in the literature (Arya et al. 2007, Li et al. 2014, Roy et al. 2019), we normalize the manufacturer's production costs to zero. We consider the retailer's online store SO to be the baseline with respect to operating costs, and normalize these costs to zero as well. Since the manufacturer often does not possess the same degree of proficiency in operating a retail channel as the retailer (Arya et al. 2007, Hotkar and Gilbert 2021), we consider the manufacturer's cost of selling to be k_O for each unit of sales through his online store O. B&M stores generally are more expensive to maintain, in part due to the cost involved in servicing and convincing potential customers to buy the product (Li et al. 2019). We consider k_S to be the cost incurred by the B&M store on each consumer that visits the store and consequently purchases the product, irrespective of the channel used for the purchase.

It is an increasingly common practice for omnichannel retailers to adopt a uniform pricing policy across their online and B&M stores (Nageswaran et al. 2020), as more than two-thirds of retailers offer identical prices across channels (Cavallo 2017). Consequently, we consider that the omnichannel retailer sets an identical retail price, p_S, across both channels, while the manufacturer sets his retail price at p_O. Figure 2 illustrates the supply chain structure. We now shift our attention to the utility, and the uncertainty therein, derived by the consumer from purchasing this product.

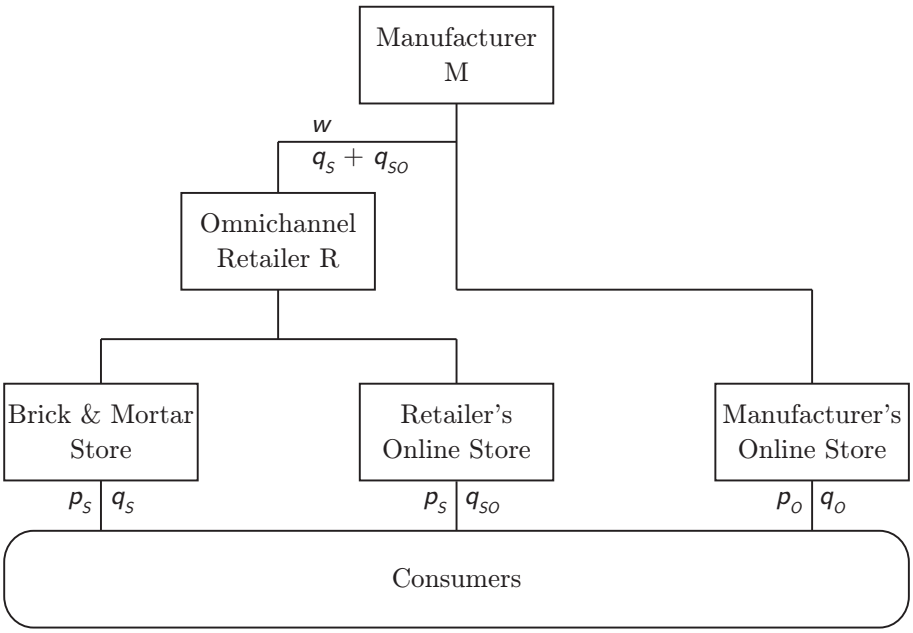


Figure 2 The Supply Chain Structure

3.2. Consumer Utility

In line with our discussion about sophisticated electronics and other high-end consumer products in Section 1, we consider a product that has significant non-digital attributes, such as a sophisticated digital device, or a high-end fashion product (Lal and Sarvary 1999, Wu et al. 2004, Balakrishnan et al. 2014, Mehra et al. 2018). Following Mehra et al. (2018), we consider that the utility of the product is v for the consumer who finds it to be her *best-fit*. Since the product has significant non-digital attributes, consumers will be able to ascertain product fitness perfectly only when they evaluate it in person. Hence, a consumer who evaluates exclusively online may purchase a product that does not provide her with the best-fit.

Let $\rho \in (0,1)$ be the probability that a consumer finds her best-fit product by exclusively evaluating online. The utility derived from purchasing the product that is not her best-fit is $v - \Delta \in (0, v)$. The reduction in utility, Δ, may arise either from accepting a product with a less-than-ideal fit, or from the hassle involved in returning a misfit product and going through the process again to find the best-fit product (Nageswaran et al. 2020). Therefore, if a consumer evaluates and buys the product exclusively online, her expected utility is $\rho v + (1-\rho)(v-\Delta)$, or $v - (1-\rho)\Delta$. Rewriting $(1-\rho)\Delta$ as δ, the expected utility for this consumer is $v - \delta < v$. We can reasonably assume parameters such as v and δ to be common knowledge to both consumers and retailers (Balakrishnan et al. 2014, Mehra et al. 2018).

Different consumers also incur different costs for choosing different channels or stores, both for evaluating the product, and subsequently buying the product. We now discuss how consumers incur these costs, and how they are heterogeneous with regard to the costs of accessing different channels.

3.3. Consumer Costs

While it is evident that evaluating and purchasing exclusively online provides consumers with a lower expected utility from the product itself, some consumers still prefer transacting entirely online. As demonstrated by recent research (Balakrishnan et al. 2014, Mehra et al. 2018, Nageswaran et al. 2020), consumers incur different costs in different channels. First, consumers incur different costs to travel to the B&M store, either to evaluate the product or to buy the product. This arises from different aspects of the consumer, such as proximity to the store, and the opportunity cost of their time spent in traveling to the store and searching therein (Balasubramanian 1998, Desai et al. 2010, Forman et al. 2009). Let us denote this travel cost by $\tau \geq 0$.

One can group the consumers into two types: (a) L-type consumers: those who incur a lower cost to visit the B&M store relative to the reduction in utility that would arise from evaluating and purchasing the product exclusive online (i.e., $\tau \leq \Delta$), and (b) H-type consumers: those for whom the cost to travel to the B&M store is greater than the reduction in utility mentioned above (i.e., $\tau > \Delta$) Coughlan and Soberman (2005), Mehra et al. (2018). Consequently, we consider that $\lambda \in (0,1)$ fraction of all consumers are L-type consumers, and $1-\lambda$ fraction are H-type consumers. In order to improve tractability, we normalize the travel cost of the L-type consumers to $\tau = 0 < \Delta$, and consider that of the H-type consumers to be $\tau = t \geq \Delta > 0$. Once the consumer visits the B&M store, we normalize the additional search and purchase costs inside the store to zero (Mehra et al. 2018).

It has been well-established that the online search process has become easier with technological improvements in e-commerce, and can be considered less expensive than the corresponding cost involved in traveling to a B&M store (Balasubramanian 1998, Forman et al. 2009, Gupta et al. 2004). Therefore, we consider that visiting online stores over the web is inexpensive, and thus consumers do not incur any cost when evaluating the product online. However, when buying online, consumers differ in their attitude toward online purchasing, and in their willingness to wait for product delivery that results in delayed gratification, (Bart et al. 2005, Frambach et al. 2007, Balakrishnan et al. 2014). We model different consumers' attitudes towards purchasing online as a uniform distribution, or the standard Hotelling line between 0 and 1, where $x \in [0,1]$ is a consumer's index proportional to her cost of purchasing from the online retailer. A consumer's position on the Hotelling line is her private information.

The proportionality constant multiplier for the online purchasing cost depends on the consumers' preference for (alternatively, level of trust in) the online retailer (Mehra et al. 2018). While in certain product categories, such as apparel, the majority of consumers prefer the online store of the manufacturer over that of the retailer (Retail Info Systems 2017), large online retailers, such as Amazon.com, are favored strongly over other e-commerce websites, including those of smaller manufacturers selling through Amazon.com (Masters 2019). Consequently, we model the online purchase

cost multiplier for the retailer's online store SO to be $c_{SO} > 0$, and that for the manufacturer's online store O to be $c_O > 0$. Depending on the product category and the reputation of the two online stores, either of these cost multipliers may be greater than the other. These cost multipliers are common knowledge to the consumers, the manufacturer, and the retailer. Let us now discuss how these utilities and costs affect a consumer's evaluating and purchasing process, and her eventual purchase decision. We present a list of the key variables and parameters of the model in Table 1.

Table 1 List of Key Variables and Parameters

Symbol	Definition
Variables	
w	Wholesale price offered by the Manufacturer
q_S	Quantity sold by the Omnichannel Retailer through her B&M Store, S
q_{SO}	Quantity sold by the Omnichannel Retailer through her Online Store, SO
$q_R = q_S + q_{SO}$	Total quantity sold by the Omnichannel Retailer through her Online Store
q_O	Quantity sold by the Manufacturer through his Online Store, O
p_S	Retail price charged by the Retailer in both her B&M and Online Stores
p_O	Retail price charged by the Manufacturer in his Online Store
Parameters	
v	Utility derived by a consumer from her best-fit product
$\rho \in (0,1)$	Probability of obtaining the best-fit product while exclusively evaluating and buying online
$\Delta \in (0, v)$	Decrease in utility from obtaining a product that is not best-fit
$\delta = \Delta(1-\rho)$	Decrease in expected utility while exclusively evaluating and buying online
$\tau \in 0, t$	Cost of traveling to the B&M Store
L	Type of consumers with low travel cost, i.e., $\tau = 0$
H	Type of consumers with high travel cost, i.e., $\tau = t$
λ	Fraction of consumers who are L-type, remaining $1-\lambda$ are H-type
$x \in [0,1]$	Heterogeneity of consumers' disutility from online purchase
$c_{SO} > 0$	Consumer's disutility multiplier for the Omnichannel Retailer's Online Store
$c_O > 0$	Consumer's disutility multiplier for the Manufacturer's Online Store
k_S	Cost incurred by the Omnichannel Retailer on every consumer visiting B&M store
k_O	Cost incurred by the Manufacturer on each unit sold through his Online Store
π_M	Profit of the Manufacturer
π_R	Profit of the Omnichannel Retailer

3.4. Consumer Purchase Decision

As consumers increasingly perform product evaluation and product purchase separately, especially for such sophisticated products (Shin 2007, Balakrishnan et al. 2014), we model the consumer's decision-making as a two-step process: first (a) evaluate the product, then (b) buy the product. Figure 3 describes the consumer evaluation and the purchase decision tree, along with the net expected utilities for the different types of consumers.

In the first step, consumers evaluate the product. Let us first focus on consumers evaluating online. In the interest of restricting focus on the practices of consumer showrooming and supplier encroachment, we remove from our consideration practices such as *webrooming* (Jing 2018) and

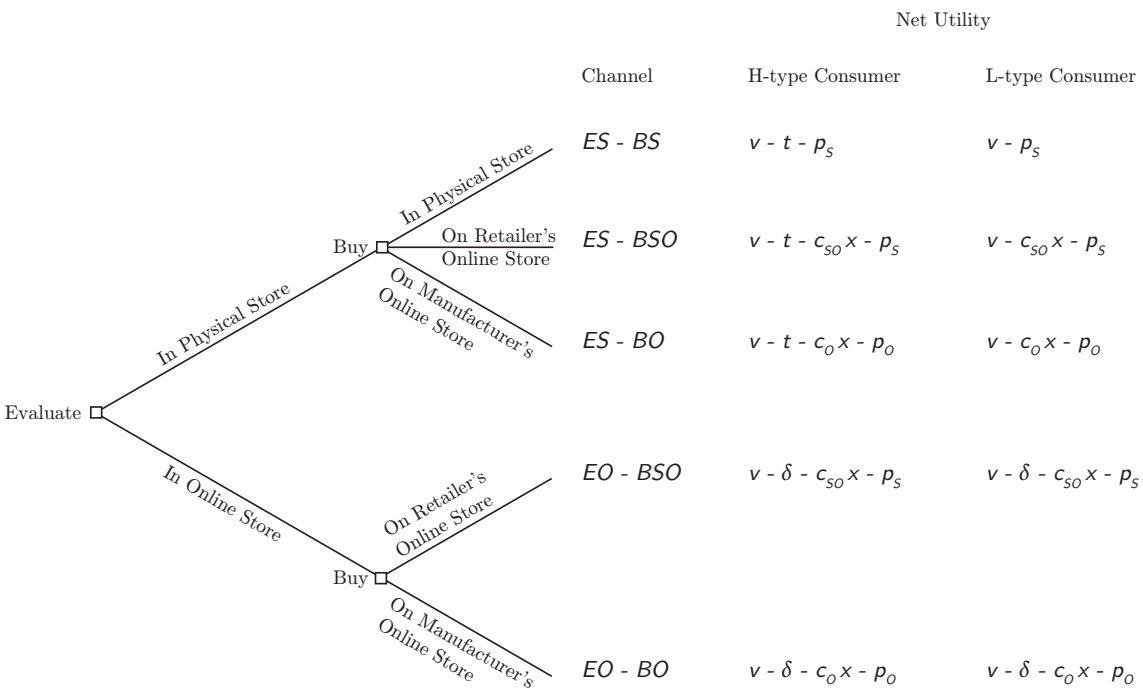


Figure 3 Consumer Evaluation and Purchase Decision Tree

buy-online-pickup-at-store (Gao and Su 2017). Therefore, we consider that any consumer, either H-type or L-type, who evaluates exclusively online, has decided not to visit the store at any step of her purchasing process, and will end up either buying from the retailer's online store SO, or from the manufacturer's online store O, or not buying at all. As described in Section 3.2, if this consumer decides to buy from either online store after evaluating exclusively online, she will derive a utility $v-\delta$. Following our discussion in Section 3.3, if she decides to buy from the retailer's online store SO, she incurs a cost of online purchase equal to $c_{SO}x$, and pays a price p_S to the retailer, resulting in a net utility of $v-\delta-c_{SO}x-p_S$. We refer to such an evaluation-purchase path as the $EO-BSO$ channel, i.e., *evaluate online, buy at retailer's online store*. Similarly, if she decides to buy from the manufacturer's online store (i.e., the $EO-BO$ channel), her net utility would be $v-\delta-c_O x-p_O$.

Let us now bring our attention to consumers evaluating in the retailer's B&M store, S. These consumers may have also read reviews online before traveling to the store, a process that we have established to be inexpensive. Therefore, by visiting the store, a consumer incurs only the travel cost τ, which is 0 for an L-type consumer, and t for an H-type consumer. By evaluating the product at the store, this consumer is guaranteed to ascertain the best-fit nature of the product (Mehra et al. 2018). If she consequently decides to buy the product at S (i.e., the $ES-BS$ channel), her net utility would be $v-\tau-p_S$. She may also decide to withhold her buying step at the B&M store, and complete the purchasing process on either online store, i.e., engage in showrooming. Since she has already found the product to be her best-fit, she incurs no further

loss of utility due to the mismatch but would incur the cost of online purchase. Therefore, by evaluating the product at the B&M store S and buying it from the retailer's online store SO, i.e., the $ES-BSO$ channel, (alternatively, the manufacturer's online store O, i.e., the $ES-BO$ channel), her net utility would be $v-\tau-c_{SO}x-p_S$ (alternatively, $v-\tau-c_Ox-p_O$).

3.5. Sequence of Events

In line with the consumer showrooming and supplier encroachment streams of literature discussed in Section 2, and the discussion with reference to the supply chain structure in Section 3.1, we model the sequence of events as a single-period Stackelberg game, where the manufacturer and the retailer engage in a wholesale price contract.

Consequently, we set up the game as follows: (i) Stage 0: The manufacturer decides whether or not to *encroach* by opening the online store O, which then becomes common knowledge; (ii) Stage 1: The manufacturer offers a wholesale price w, to which the retailer responds with her purchase quantity $q_R=q_S+q_{SO}$ for her B&M and online stores S and SO; (iii) Stage 2: The manufacturer and the retailer simultaneously determine their selling prices p_S and p_O; (iv) Stage 3: Consumers observe the retail prices and decide which store to use (if any) for evaluation and for purchase based on their private cost information as well as the common information set, and the market clears.

Figure 4 illustrates the sequence of events in the game, based on which we next model the optimization problems for the manufacturer and the retailer.

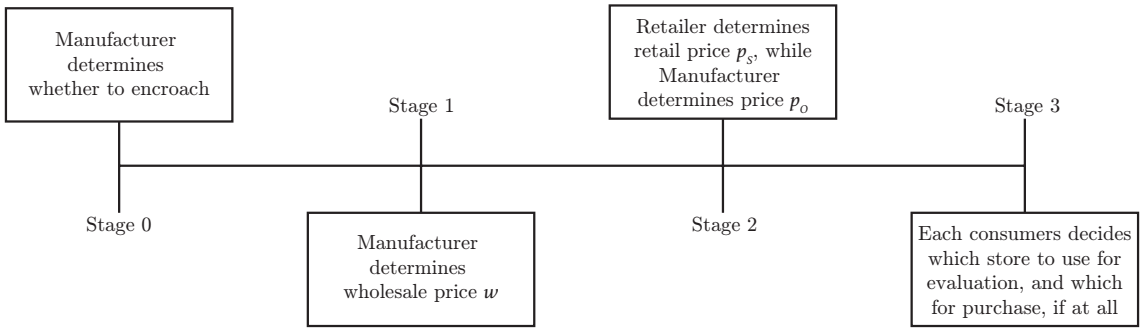


Figure 4 Sequence of Events

3.6. Optimization Problems for the Manufacturer and the Retailer

The manufacturer chooses the wholesale price w in Stage 1 and the retail price p_O in Stage 2, in order to maximize his total profit, π_M, which comprises of his profit from the wholesale channel, π_{M1}, and his profit from the retail channel, π_{M2}, given by:

$$\begin{aligned}&\pi_M=\pi_{M1}+\pi_{M2},\ \text{where}\\&\pi_{M1}=wq_R=w\left(q_S+q_{SO}\right),\ \text{and}\ \pi_{M2}=\left(p_O-k_O\right)q_O,\end{aligned}\tag{1}$$

where q_S, q_{SO}, and q_O are the quantities sold through the retailer's B&M channel, her online channel, and the manufacturer's online channel, respectively, and k_O is the cost incurred by the manufacturer on each unit sold.

On the other hand, the retailer chooses her purchase quantities q_S (for B&M store) and q_{SO} (for online store) in response to the manufacturer's offer w from Stage 1, and retail prices p_S and p_{SO} in Stage 2, in order to maximize her total profit, π_R, which comprises of her profit from the B&M store, π_{RS}, and her profit from the retail channel, π_{RSO}, given by:

$$\begin{aligned} &\pi_R = \pi_{RS} + \pi_{RSO}, \text{ where} \\ &\pi_{RS} = (p_S - w)\, q_S - k_S n S, \text{ and } \pi_{RSO} = (p_S - w)\, q_{SO}, \end{aligned} \tag{2}$$

where k_S is the cost incurred by the retailer to serve every visitor to the B&M store, and n_S is the total number of visitors to the B&M store. These visitors include consumers who purchase at the B&M store and consumers who showroom.

In Stage 3, H-type and L-type consumers observe retail prices p_S, p_{SO}, and p_O, before determining where to evaluate the product, and if buying, through which channel to buy the product. Let us denote the fraction of H-type consumers who buy from the S, SO, and O stores as q_{SH}, q_{SOH}, and q_{OH}, respectively. Similarly, the corresponding fractions for the L-type consumers are denoted by q_{SL}, q_{SOL}, and q_{OL}, respectively. Since we consider that the L-type (conversely, H-type) consumers constitute λ (conversely, $1-\lambda$) proportion of the consumer population, the total quantities sold in the three channels, for a given set of retail prices p_S and p_O, are given by

$$q_S(p_S, p_O) = (1-\lambda)\left[q_{SH}(p_S, p_O)\right]^+ + \lambda\left[q_{SL}(p_S, p_O)\right]^+, \tag{3a}$$

$$q_{SO}(p_S, p_O) = (1-\lambda)\left[q_{SOH}(p_S, p_O)\right]^+ + \lambda\left[q_{SOL}(p_S, p_O)\right]^+, \text{ and} \tag{3b}$$

$$q_O(p_S, p_O) = (1-\lambda)\left[q_{OH}(p_S, p_O)\right]^+ + \lambda\left[q_{OL}(p_S, p_O)\right]^+, \tag{3c}$$

where the superscript $+$ denotes non-negative quantities.

In stage 2, after the manufacturer has announced wholesale price w, the retailer and the manufacturer simultaneously set their retail prices p_S and p_O, in order to maximize their stage 2 profits, given by π_{M2} in (1) and π_R in (2), respectively. Substituting (3a) through (3c) in (1) and (2), the stage 2 optimization problems for the manufacturer and the retailer become:

$$\begin{aligned} \underset{p_O \geq 0}{\text{maximize}}\ \pi_{M2} \quad &= (p_O - k_O)\left((1-\lambda)\left[q_{OH}(p_S, p_O)\right]^+ + \lambda\left[q_{OL}(p_S, p_O)\right]^+\right), \\ &\text{s. t. } q_{SOH}, q_{SOL} \geq 0, \text{ and} \end{aligned} \tag{4}$$

$$\begin{aligned} \underset{p_S \geq 0}{\text{maximize}}\ \pi_R \quad &= (p_S - w)\left((1-\lambda)\left[q_{SH}(p_S, p_O)\right]^+ + \lambda\left[q_{SL}(p_S, p_O)\right]^+\right) \\ &\quad + (p_S - w)\left((1-\lambda)\left[q_{SOH}(p_S, p_O)\right]^+ + \lambda\left[q_{SOL}(p_S, p_O)\right]^+\right) - k_S n S, \\ &\text{s. t. } q_{SH}, q_{SL}, q_{OH}, q_{OL} \geq 0. \end{aligned} \tag{5}$$

Simultaneously optimizing (4) for the manufacturer and (5) for the retailer in stage 2, we obtain, as a function of the wholesale price w offered by the manufacturer in stage 1, the optimal retail prices $p_S^*(w)$, $p_{SO}^*(w)$, and $p_O^*(w)$. Substituting these retail prices in (3a) through (3c), we obtain the optimal selling quantities as a function of w, $q_S^*(w)$, $q_{SO}^*(w)$, and $q_O^*(w)$.

In stage 1, the manufacturer anticipates these responses to his wholesale price offer. Substituting $q_S^*(w)$, $q_{SO}^*(w)$, and $q_O^*(w)$, obtained from the optimization in stage 2, in (1), we can rewrite the manufacturer's optimization problem in stage 1:

$$\begin{aligned} \underset{w\geq 0}{\text{maximize}}\ \pi_M =& w\left(q_S^*(w)+q_{SO}^*(w)\right)+\left(p_O^*(w)-k_O\right)q_O^*(w), \\ &\text{s. t. } q_S^*(w), q_{SO}^*(w), q_O^*(w)\geq 0. \end{aligned} \tag{6}$$

The optimization problem in (6) yields the equilibrium wholesale price, w^*, that the manufacturer will offer. Substituting this w^* in $q_S^*(w)$, $q_{SO}^*(w)$, and $q_O^*(w)$, obtained from stage 2, we arrive at the equilibrium q_S^*, q_{SO}^*, and q_O^*, and the corresponding equilibrium selling prices, p_S^*, p_{SO}^*, and p_O^*. Through both of her channels, the retailer sells the total quantity $q_R^* = q_S^* + q_{SO}^*$. Finally, substituting these equilibrium quantities and prices in their respective profit functions given by (1) and (2), we determine the equilibrium profits for the manufacturer and the retailer.

4. Model Analysis

In this section, we present analysis of the model and the main results, derived using the steps outlined in Section 3.6. For the purpose of easing the exposition and maintaining analytical tractability, we normalize some of the variables described in Section 3. As we discuss in Section 4.1, these simplifications do not alter the fundamental nature of the model, nor do they take away the qualitative essence of the insights; instead, they allow us to focus on the important strategic interactions.

4.1. Normalization of Variables

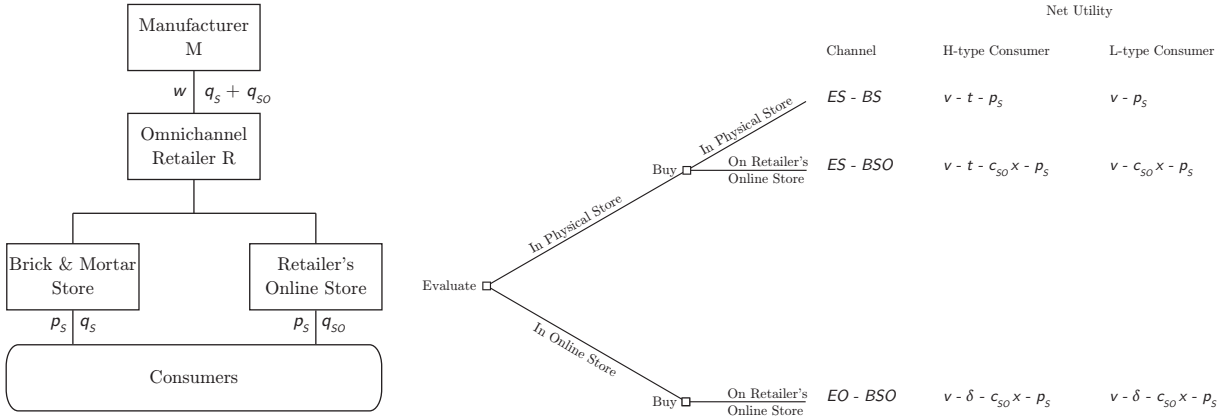
Without loss of generality, we normalize v, the utility derived by consumers from consuming one unit of the product to 1. The results hold qualitatively if v is scaled to any arbitrarily large number. We then consider the expected loss of utility from misfits resulting from exclusively evaluating and purchasing online, denoted by δ. We consider $\delta = \frac{1}{2}$, which satisfies the logical requirement that $0 \leq \delta \leq 1$, and the results qualitatively hold for any $0 \leq \delta \leq 1$.

Recall that for the H-type consumers, $t > \Delta$, which ensures that not all consumers end up showrooming. Since $\delta = (1-\rho)\Delta$, and $0 < \rho < 1$, it is evident that $t > \delta$. Further, we want to consider the strategic tension between the three channels, and do not want to systematically eliminate the H-type consumers coming to the B&M store even at $p_S = 0$. This results in an upper bound of t, namely $t < v = 1$. Consequently, we consider $\frac{1}{2} < t < 1$. Next, let us consider c_{SO} and

c_O, the disutility multiplier faced by a consumer when buying from the omnichannel retailer's online store, and the manufacturer's online store, respectively. By normalizing $c_{SO} = 1$, we are able to analyze two distinct scenarios in an elegant form: (a) $c_O \leq 1$ and (b) $c_O > 1$, and obtain clean managerial and theoretical insights.

4.2. Baseline: The No-encroachment Scenario

In order to understand how supplier encroachment and consumer showrooming affect the relationship between the manufacturer and the omnichannel retailer, we first discuss the dynamics between the two channel partners when the manufacturer does not encroach, that is, there exists no direct channel. While this baseline has been discussed in the supply chain management and encroachment literature extensively (Arya et al. 2007, Anand et al. 2008, Li et al. 2014, Hotkar and Gilbert 2021, Roy et al. 2019), it is also important to realize an important distinguishing factor in our model. Unlike most of the encroachment literature, we explicitly model the consumer's purchasing behavior in the no-encroachment baseline, as we also do in the model where the supplier encroaches, making use of the heterogeneity in net consumer utilities. Further, we treat the baseline as an exogenous condition imposed on the manufacturer, such that the manufacturer is unable to encroach. For this baseline solution to be an equilibrium solution, the manufacturer needs to credibly communicate his inability to encroach, which we discuss in Section 5.



(a) Supply Chain Structure (b) Consumer Decision Tree

Figure 5 Supply Chain Structure and Consumer Decision Tree under No-encroachment

When the manufacturer does not encroach, his online store O is eliminated from the supply chain structure, and from the consumer's evaluation and purchase decision tree. Consequently, Figure 2 and Figure 3 transform to Figure 5a and Figure 5b, respectively. Let us now analyze the net utilities received by the L-type and H-type consumers. Since $\delta, c_{SO} > 0$, for the L-type consumer, the

utility from the $ES-BS$ channel dominates those from the $ES-BSO$ and $EO-BSO$ channels. Further, since $t>\delta$, for the H-type consumer, the $ES-BSO$ channel is strictly dominated by the $EO-BSO$ channel. The H-type consumer indifferent between the $EO-BSO$ and the $ES-BS$ channels is given by $x=\frac{t-\delta}{c_{SO}}$, and she who is indifferent between the $EO-BSO$ channel and not buying at all is given by $x=\frac{v-\delta-p_S}{c_{SO}}$. Observe that in the absence of supplier encroachment, neither type of consumer showrooms: the purchase happens in the same channel as the evaluation.

Let us now apply the normalization rules described in Section 4.1, i.e., $v=1$, $\delta=\frac{1}{2}$, $c_{SO}=1$, and $\frac{1}{2}<t<1$. Given retail price p_S determined in stage 2, the demand functions for the L and H consumer types realized in stage 3 are given by (7a) and (7b), respectively:

$$\left(q_{SL}^{NE}(p_S),q_{SOL}^{NE}(p_S)\right)=\begin{cases}(1,0), & \text{if } p_S\le 1,\\ (0,0), & \text{otherwise, and,}\end{cases} \tag{7a}$$

$$\left(q_{SH}^{NE}(p_S),q_{SOH}^{NE}(p_S)\right)=\begin{cases}\left(\frac{3}{2}-t,t-\frac{1}{2}\right), & \text{if } p_S\le 1-t,\\ \left(0,\frac{1}{2}-p_S\right), & \text{if } p_S\in\left(1-t,\frac{1}{2}\right],\\ (0,0), & \text{otherwise,}\end{cases} \tag{7b}$$

where q_{SL} (alternatively, q_{SH}) and q_{SOL} (alternatively, q_{SOH}) denote the fraction of L-type (alternatively, H-type) consumers who buy from the B&M (S) and retailer's online (SO) stores, respectively. Superscript NE refers to the *no encroachment* scenario. Substituting (7a) and (7b) in (3a) and (3b), we obtain the retailer's selling quantities in the S and SO stores, as a function of the retail price p_S, given by (8a). Using (8a), we can obtain her total selling quantity across the two channels, $q_R^{NE}=q_S^{NE}+q_{SO}^{NE}$, as a function of the retail price p_S. The number of consumers serviced by the B&M store is equal to the number of consumers buying from the B&M store since there is no showrooming. Therefore, $n_S^{NE}(p_S)=q_S^{NE}(p_S)$. (8b) provides $q_R^{NE}(p_S)$ and $n_S^{NE}(p_S)$.

$$\left(q_S^{NE}(p_S),q_{SO}^{NE}(p_S)\right)=\begin{cases}\left(\left(\frac{3}{2}-t\right)(1-\lambda)+\lambda,\left(t-\frac{1}{2}\right)(1-\lambda)\right), & \text{if } p_S\le 1-t,\\ \left(\lambda,\left(\frac{1}{2}-p_S\right)(1-\lambda)\right), & \text{if } p_S\in\left(1-t,\frac{1}{2}\right],\\ (\lambda,0), & \text{if } p_S\in\left(\frac{1}{2},1\right],\\ (0,0), & \text{otherwise, and,}\end{cases} \tag{8a}$$

$$\left(q_R^{NE}(p_S),n_S^{NE}(p_S)\right)=\begin{cases}\left(1,\left(\frac{3}{2}-t\right)(1-\lambda)+\lambda\right), & \text{if } p_S\le 1-t,\\ \left(\frac{1+\lambda}{2}-p_S(1-\lambda),\lambda\right), & \text{if } p_S\in\left(1-t,\frac{1}{2}\right],\\ (\lambda,\lambda), & \text{if } p_S\in\left(\frac{1}{2},1\right],\\ (0,0), & \text{otherwise.}\end{cases} \tag{8b}$$

In stage 2, the only optimization problem is that of the retailer, as there is no direct channel for the manufacturer. Substituting (8b) in the retailer's optimization problem in (5), we obtain the retailer's stage 2 optimization problem under no-encroachment:

$$\underset{p_S\ge 0}{\text{maximize}}\ \pi_R^{NE}=\begin{cases}(p_S-w)-k_S\left(\left(\frac{3}{2}-t\right)(1-\lambda)+\lambda\right), & \text{if } p_S\le 1-t,\\ (p_S-w)\left(\frac{1+\lambda}{2}-p_S(1-\lambda)\right)-k_S\lambda, & \text{if } p_S\in\left(1-t,\frac{1}{2}\right],\\ (p_S-w)\lambda-k_S\lambda, & \text{if } p_S\in\left(\frac{1}{2},1\right],\\ 0, & \text{otherwise.}\end{cases} \tag{9}$$

4.3. Supplier Encroachment Scenario

Let us now analyze the net utilities received by the L-type and H-type consumers. Since $c_O, c_{SO} > 0$, for the L-type consumer, the utility from the $ES-BS$ and $ES-BO$ channels dominate those from the $ES-BSO$, $EO-BSO$ and $EO-BO$ channels. The consumer indifferent between these two channels is given by $x = \frac{p_S - p_O}{c_O}$. Further, since $t > \delta$, for the H-type consumer, the $ES-BSO$ channel is strictly dominated by the $EO-BSO$ channel, and the $ES-BO$ channel is strictly dominated by the $EO-BO$ channel. The H-type consumer indifferent between the $EO-BSO$ and the $ES-BS$ channels is given by $x = \frac{t-\delta}{c_{SO}}$, she who is indifferent between $ES-BS$ and $EO-BSO$ channels is given by $x = \frac{t-\delta+p_S-p_O}{c_O}$, and who is indifferent between $EO-BO$ and $EO-BSO$ channels is given by $x = \frac{p_O-p_S}{c_{SO}-c_O}$. Next, we apply the normalization rules and determine the demand for each channel as below:

$$q_S(p_S, p_O) = 1 + \frac{(p_O - p_S)\lambda}{c_O} + \left(t - \frac{1}{2}\right)(\lambda - 1) \tag{10a}$$

$$q_{SO}(p_S, p_O) = \left(t - \frac{1}{2} - \frac{p_S - p_O}{1 - c_O}\right)(1 - \lambda) \tag{10b}$$

$$q_O(p_S, p_O) = \frac{(p_O - p_S)(c_O - \lambda)}{c_O(1 - c_O)} \tag{10c}$$

Substituting the above demands in (4) and (5), we get the optimization problem for the retailer and the manufacturer in Stage 2. We obtain the optimal retail prices and quantities as a function of the wholesale price w as below:

$$p_O^*(w) = \frac{c_O - c_O(c_O + 2k_O + w) + (2k_O + w)\lambda}{3(\lambda - c_O)} \tag{11a}$$

$$p_S^*(w) = \frac{1}{3}\left(k_O + 2w + \frac{2(1 - c_O)c_O}{\lambda - c_O}\right) \tag{11b}$$

$$q_O^*(w) = \frac{c_O(c_O - 1 - k_O + w) + (k_O - w)\lambda}{3(c_O - 1)c_O} \tag{11c}$$

$$q_S^*(w) = \frac{1}{6}\left(7 + 2c_O - 6t - 3\lambda + \frac{2(k_O + 3c_O t - w)\lambda}{c_O} + \frac{2(c_O - 1)c_O}{\lambda - c_O}\right) \tag{11d}$$

$$q_{SO}^*(w) = \frac{(\lambda - 1)(c_O(c_O(5 - 6t) - 2k_O - 5 + 6t + 2w) + (3 + 2k_O - 6t + c_O(6t - 3) - 2w)\lambda)}{6(c_O - 1)(c_O - \lambda)} \tag{11e}$$

We place the above optimal price and quantity in the stage 1 optimization problem (6) and derive the equilibrium values for this market condition as:

$$w^* = \frac{k_O}{4}\frac{2(1 - c_O)c_O}{(\lambda - c_O)} \tag{12a}$$

$$q_O^* = \frac{4c_O - 4c_O^2 + c_O k_O - k_O \lambda}{(4c_O - 4c_O^2)} \tag{12b}$$

$$q_S^* = \frac{1}{4}\left(4\left(\frac{1}{2} + c_O - t\right) + \frac{4(1 - c_O)c_O}{(c_O - \lambda)} + \left(-2 + \frac{k_O}{c_O} + 4t\right)\lambda\right) \tag{12c}$$

$$q_{SO}^* = \frac{(1-\lambda)\left(4c_O^2\left(\frac{1}{2}-t\right)-4c_O+(2+k_O-4t)\lambda+c_O\left(4c_O-k_O+4\left(t-\frac{1}{2}\right)(1+\lambda)\right)\right)}{4(1-c_O)(\lambda+c_O)} \tag{12d}$$

$$p_O^* = \frac{3k_O}{4}+\frac{(1-c_O)c_O}{\lambda-c_O} \tag{12e}$$

$$p_S^* = \frac{k_O}{2}+\frac{2(1-c_O)c_O}{\lambda-c_O} \tag{12f}$$

We now utilize the above results to highlight the effect of a critical parameter - the manufacturer's operating cost (i.e., k_O) - that has been the focus of the literature on supplier encroachment (Arya et al. 2007, Li et al. 2014, Hotkar and Gilbert 2021).

LEMMA 1. *When the operating cost (k_O) for the manufacturer's online store increases,*

(a) manufacturer's online retail price (p_O) increases; retailer's price (p_S) also increases, but at a lower rate than that of the manufacturer.

(b) demand in the manufacturer's online store (q_O) goes down. On the other hand, since the market is covered, the demand increases for the retailer (q_R).

(c) item wholesale price (w) for the omnichannel retailer increases.

Proof of Lemma 1. Differentiating w^*, q_O^*, q_S^*, q_{SO}^*, p_O^*, and p_S^* from (12a) through (12f) with respect to k_O, we find that

(a) $\frac{\partial p_O^*}{k_O} > 0$ and $\frac{\partial p_S^*}{k_O} > 0$ $\frac{\partial p_O^*}{k_O} > \frac{\partial p_S^*}{k_O}$;

(b) $\frac{\partial q_O^*}{k_O} < 0$, $\frac{\partial q_S^*}{k_O} > 0$, and $\frac{\partial q_{SO}^*}{k_O} > 0$;

(c) $\frac{\partial w^*}{k_O} > 0$. □

Since the manufacturer's online store operation becomes expensive (k_O), the manufacturer increases the retail price at his online store. This eventually leads to lower demand at the online store (q_O). Since we consider our market covered, the lost demand at the manufacturer goes to the omnichannel retailer (q_R). The retailer, then, takes advantage of the increasing demands at her store and increases the retail price at her store (p_S). On the other hand, the manufacturer exploits the increased demand at the retailer's stores and increases the wholesale price (w) offered to the retailers in order to make up for the lost business in the direct channel. Using this result, we develop and discuss some of our key findings in the next section.

5. Key Findings and Managerial Insights

In this section, we focus on how changing different parameters impact the online manufacturer and the omnichannel retailer. The operating cost of the manufacturer's direct channel plays an essential role in determining whether the manufacturer should set up the direct channel. From the literature, we observe that there exist some upper threshold limits for the manufacturer's online channel operating costs beyond which the manufacturer makes no profit, and hence, in that case, the

manufacturer does not encroach (Arya et al. 2007, Li et al. 2014, Hotkar and Gilbert 2021). In the analysis below, we observe how the cost of the manufacturer's online store operation impacts the retailer as well as the manufacturer when encroachment and showrooming occur simultaneously.

PROPOSITION 1. *At a high cost of traveling τ, when the cost of operation at the manufacturer's online store increases (k_O), the profit of both the manufacturer and the retailer decreases; similarly, at a low traveling cost, when k_O increases, the profit for both increases.*

Proof of Proposition 1. Similar to the Proof of Lemma 1. □

Consumers cease to visit the B&M store for evaluating or buying because of high τ. So, they prefer to buy from online channels. The profit earned from purchases in the retailer's B&M store ceases and hence, the profit of the retailer goes down. When the cost incurred in maintaining the online store by the manufacturer is low, the manufacturer can attract more consumers with lower retail price p_O and earn better profit. As k_O increases, p_O also increases, and the manufacturer loses consumers, eventually leading to lower profit. When the travel cost is low, consumers prefer to visit the B&M store, and hence, the purchases happen at both the channels of the retailer, thus increasing the retailer's profit. For the manufacturer, as the value of k_O increases, he increases p_O and loses consumers to the retailer (since the market is covered). The retailer's demand increases, which again contributes to higher profits for the retailer. The manufacturer makes up for the lost business in the direct channel by increasing the wholesale price. With increased demand and higher profits, the retailer can afford the increased wholesale price set by the manufacturer, thus ultimately contributing to increasing the manufacturer's profit under these conditions.

PROPOSITION 2. *When the cost of traveling is high and the cost of operation at the manufacturer's online store is low, the manufacturer earns higher profit than the retailer; similarly, when the cost of traveling is low, and the manufacturer's cost of the store operation is high, the retailer earns more profit than the manufacturer.*

Proof of Proposition 2. Similar to the Proof of Lemma 1. □

High traveling cost implies that consumers will prefer buying from online channels. The consumers who like to visit the store to evaluate or to buy will cease to do so. This implies that the profit through the B&M channel of the retailer will lower significantly. Furthermore, low operating costs at the manufacturer's online store will enable him to keep the retail price low, attracting a higher volume of demand. Hence, the manufacturer earns more profit than the retailer in this case. On the contrary, when travel cost is low, consumers prefer to visit the B&M retail store to evaluate and buy products. The high operating cost at the manufacturer's online store means higher retail price and hence, the manufacturer will lose demand. Since the market is covered, the omnichannel

retailer will get this demand over her two channels. Therefore, when the travel cost is low and the manufacturer has a higher operating cost, the retailer will earn more profit than the manufacturer.

These insights will be beneficial for managers from both the retailer and manufacturer to run their operations strategically. From Propositions 1 and 2, we observe four scenarios arise when an upstream manufacturer with an online store encroaches the downstream omnichannel retailer's market: a) when traveling cost τ is high and the cost of operation at manufacturer's online store k_O increases, the profits of both the manufacturer and the retailer is diminished; b) when τ is low and k_O increases, both retailer's and manufacturer's profit grow; c) when τ is high and the manufacturer's online store operation cost k_O is low, the manufacturer earns higher profit than that of the retailer; d) when τ is low, and k_O is high, the retailer earns more profit than the manufacturer. One can conclude that cases (a) and (c) are not favorable for the retailer and one primary reason for that is the high traveling cost. One cause of such high traveling costs for the retailer's physical stores may be that the retailer's network of B&M stores is very sparse in a particular locality. As a result, the consumers in that locality have to endure a high traveling cost to the physical store and hence, prefer to buy online. In that case, the managers and policymakers at the retailer may attempt to increase the retailer's network density by setting up B&M stores in more frequent locations. This will lower the traveling cost and in that case, will benefit both the retailer's and the manufacturer's profits. Another way the retailer may improve her business in scenarios (a) and (c) is bundling the product coming from the upstream manufacturer with a related product and selling them as a single combined unit in both the channels – the B&M store and the online store. For example, Best Buy decided to bundle together Google Pixel phone and screen protector together as a single combined unit, a deal consumers won't find in the online Google store.

In scenario (d), the retailer earns more profit than the manufacturer. In that case, the managers at the manufacturer's online store may make an effort to reduce the cost of operation k_O at the online store. The operating cost of an online store has a variable component (for example, website troubleshooting cost or digital market cost) and a fixed component (for example, periodic maintenance and employee cost). The manufacturer should come up with a strategy to reduce at least one of the operation cost components in scenario (d). A low k_O will be beneficial for the manufacturer to be better off than the retailer.

6. Conclusion and Discussion

Our paper investigates a scenario where supplier encroachment and consumer showrooming happen in parallel. We consider a scenario where an upstream supplier establishes an online channel to encroach the downstream omnichannel retailer's market, while encouraging certain consumers to "showroom", or "browse-and-switch". We take into account two types of consumers:

(a) L-type consumers: those who incur a lower cost to visit the B&M store relative to the reduction in utility that would arise from evaluating and purchasing the product exclusively online, and (b) H-type consumers: those for whom the cost to travel to the B&M store is greater than the reduction in utility mentioned above.

We develop a model where the manufacturer offers the wholesale price and the retailer buys products for both of her channels – the B&M store and online channel, in the first stage. In the next stage, both the manufacturer and the retailer establish their retail prices across their respective channels. Finally, consumers observe the retail prices and decide which store to use (if any) for evaluation and for purchase.

We solve the model by backward induction and find the equilibrium values for the market condition where all the channels players and all the channels are present. We derive key insights from this research that may be useful for industry managers. We find that profits of both manufacturer and retailer decrease with increasing manufacturer's store operation cost when there is a high cost of traveling involved; however, when traveling cost is low, as the cost of operation increases, profits of both increase. Another result we derive from this research is that with a high traveling cost and low cost of the manufacturer's store operation, the manufacturer earns more profit than the retailer; when the traveling cost is low and the cost of operation is high, the retailer's profit is more than that of the manufacturer. In the conditions where the manufacturer is better off, the retailer may attempt to increase her physical store network density to reduce the high cost of traveling or employ strategies such as product bundling or exclusive store brand to attract more consumers to the physical as well as online stores. In unfavorable cases, the manufacturer should work on decreasing the cost of his online store operations in order to be better off.

Future extensions for this stream of research may include the following. First, it will be of interest to academics and practitioners alike to find the equilibrium conditions for the market scenario where all three players are not present and study what are the conditions that benefit the online manufacturer and what conditions benefit the omnichannel retailer. Moreover, we would like to explore how the dynamics will change if the retailer has an exclusive brand of products. It is a common strategy employed by many retailers to have their own product line in the shop called the store brand. It will be interesting to see how the market scenario will change in response to such a strategy adopted by the retailer.

References

Anand K, Anupindi R, Bassok Y (2008) Strategic inventories in vertical contracts. *Management Science* 54(10):1792–1804.

Antia KD, Bergen M, Dutta S (2004) Competing with gray markets. *MIT Sloan Management Review* 46(1):63.

Arya A, Mittendorf B (2018) Bricks-and-mortar entry by online retailers in the presence of consumer sales taxes. *Management Science* 64(11):5220–5233.

Arya A, Mittendorf B, Sappington DE (2007) The bright side of supplier encroachment. *Marketing Science* 26(5):651–659.

Balakrishnan A, Sundaresan S, Zhang B (2014) Browse-and-switch: Retail-online competition under value uncertainty. *Production and Operations Management* 23(7):1129–1145.

Balasubramanian S (1998) Mail versus mall: A strategic analysis of competition between direct marketers and conventional retailers. *Marketing Science* 17(3):181–195.

Bart Y, Shankar V, Sultan F, Urban GL (2005) Are the drivers and role of online trust the same for all web sites and consumers? a large-scale exploratory empirical study. *Journal of marketing* 69(4):133–152.

Bell DR, Gallino S, Moreno A (2018) Offline showrooms in omnichannel retail: Demand and operational benefits. *Management Science* 64(4):1629–1651.

Cachon GP (2003) Supply chain coordination with contracts. *Handbooks in operations research and management science* 11:227–339.

Cain S (2018) 'I'd love to scream at them': How showroomers became the no 1 threat to bookshops. https://www.theguardian.com/books/shortcuts/2018/dec/17/id-love-to-scream-at-them-how-showroomers-became-the-no-1-threat-to-bookshops, (Accessed on 06/24/2020).

Cao S (2019) How shoppers search online while in-store - think with google. https://www.thinkwithgoogle.com/consumer-insights/online-in-store-shopping-search/, (Accessed on 06/26/2020).

Carlton DW, Chevalier JA (2001) Free riding and sales strategies for the internet. *The Journal of Industrial Economics* 49(4):441–461.

Caro F, Kök AG, Martínez-de Albéniz V (2020) The future of retail operations. *Manufacturing & Service Operations Management* 22(1):47–58.

Cattani K, Gilland W, Heese HS, Swaminathan J (2006) Boiling frogs: Pricing strategies for a manufacturer adding a direct channel that competes with the traditional channel. *Production and Operations Management* 15(1):40.

Cavallo A (2017) Are online and offline prices similar? evidence from large multi-channel retailers. *American Economic Review* 107(1):283–303.

Chiang WyK, Chhajed D, Hess JD (2003) Direct marketing, indirect profits: A strategic analysis of dual-channel supply-chain design. *Management science* 49(1):1–20.

Coughlan AT, Soberman DA (2005) Strategic segmentation using outlet malls. *International Journal of Research in Marketing* 22(1):61–86.

Cui TH, Kong G, Pourghannad B (2020) Is simplicity the ultimate sophistication? the superiority of linear pricing. *Production and Operations Management* .

Davis D (2019) Consumer brands make big bets on ecommerce. https://www.digitalcommerce360.com/2019/05/28/consumer-brands-make-big-bets-on-ecommerce/, (Accessed on 06/25/2020).

Desai PS, Krishnamoorthy A, Sainam P (2010) "call for prices": Strategic implications of raising consumers' costs. *Marketing Science* 29(1):158–174.

Dumrongsiri A, Fan M, Jain A, Moinzadeh K (2008) A supply chain model with direct and retail channels. *European Journal of Operational Research* 187(3):691–718.

Forman C, Ghose A, Goldfarb A (2009) Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management science* 55(1):47–57.

Frambach RT, Roest HC, Krishnan TV (2007) The impact of consumer internet experience on channel preference and usage intentions across the different stages of the buying process. *Journal of interactive marketing* 21(2):26–41.

Gao F, Su X (2017) Omnichannel retail operations with buy-online-and-pick-up-in-store. *Management Science* 63(8):2478–2492.

Google Brand Store (2020) Google brand store: Google products - best buy. https://www.bestbuy.com/site/brands/google/pcmcat344500050000.c, (Accessed on 06/25/2020).

Google Store (2020) Google store for google made devices & accessories. https://store.google.com/us/?hl=en-US&&43700028976493334, (Accessed on 06/25/2020).

Guan H, Gurnani H, Geng X, Luo Y (2019) Strategic inventory and supplier encroachment. *Manufacturing & Service Operations Management* 21(3):536–555.

Gupta A, Su BC, Walter Z (2004) An empirical study of consumer switching from traditional to electronic channels: A purchase-decision process perspective. *International Journal of Electronic Commerce* 8(3):131–161.

Ha A, Long X, Nasiry J (2016) Quality in supply chain encroachment. *Manufacturing & Service Operations Management* 18(2):280–298.

Hollister S (2020) Google now has a 'multibillion-dollar' hardware business. https://www.theverge.com/2020/2/3/21121492/google-hardware-multi-billion-dollar-business-q4-2019-earnings, (Accessed on 06/25/2020).

Hong Y, Pavlou PA (2014) Product fit uncertainty in online markets: Nature, effects, and antecedents. *Information Systems Research* 25(2):328–344.

Hotkar P, Gilbert SM (2021) Supplier encroachment in a nonexclusive reselling channel. *Management Science* 67(9):5821–5837.

Huang S, Guan X, Chen YJ (2018) Retailer information sharing with supplier encroachment. *Production and Operations Management* 27(6):1133–1147.

Jing B (2018) Showrooming and webrooming: Information externalities between online and offline sellers. *Marketing Science* 37(3):469–483.

Keyes D (2020) E-commerce topped 10% of us retail in 2019. https://www.businessinsider.com/ecommerce-topped-10-percent-of-us-retail-in-2019-2020-2, (Accessed on 06/24/2020).

Lal R, Sarvary M (1999) When and how is the internet likely to decrease price competition? *Marketing Science* 18(4):485–503.

Li G, Li L, Sun J (2019) Pricing and service effort strategy in a dual-channel supply chain with showrooming effect. *Transportation Research Part E: Logistics and Transportation Review* 126:32–48.

Li Z, Gilbert SM, Lai G (2014) Supplier encroachment under asymmetric information. *Management science* 60(2):449–462.

Li Z, Gilbert SM, Lai G (2015) Supplier encroachment as an enhancement or a hindrance to nonlinear pricing. *Production and Operations Management* 24(1):89–109.

Masters K (2019) 89% of consumers are more likely to buy products from amazon than other e-commerce sites: Study. https://www.forbes.com/sites/kirimasters/2019/03/20/study-89-of-consumers-are-more-likely-to-buy-products-from-amazon-than-other-e-commerce-sites/#6e7d69044af1, (Accessed on 06/28/2020).

Mehra A, Kumar S, Raju JS (2018) Competitive strategies for brick-and-mortar stores to counter "showrooming". *Management Science* 64(7):3076–3090.

Moon B (2016) The google shop at best buy: get hands on with google gear. https://blog.bestbuy.ca/smartphones-accessories/the-google-shop-at-best-buy-the-best-way-to-get-hands-on-with-google-gear, (Accessed on 06/25/2020).

Nageswaran L, Cho SH, Scheller-Wolf A (2020) Consumer return policies in omnichannel operations. *Management Science* .

O'Connell L (2020) Total retail sales in the u.s. 2012-2024. https://www.statista.com/statistics/443495/total-us-retail-sales/, (Accessed on 06/24/2020).

Pearson B (2017) Whatever happened to showrooming? how big retail is taking back the shopping trip. https://www.forbes.com/sites/bryanpearson/2017/08/16/whatever-happened-to-showrooming-how-big-retail-is-taking-back-the-shopping-trip/#72bf828879cb, (Accessed on 06/24/2020).

Retail Info Systems (2017) Why are shoppers choosing brand sites over retailers? https://risnews.com/why-are-shoppers-choosing-brand-sites-over-retailers, (Accessed on 06/28/2020).

Roy A, Gilbert SM, Lai G (2019) The implications of visibility on the use of strategic inventory in a supply chain. *Management Science* 65(4):1752–1767.

Shin J (2007) How does free riding on customer service affect competition? *Marketing Science* 26(4):488.

Singley RB, Williams MR (1995) Free riding in retail stores: An investigation of its perceived prevalence and costs. *Journal of Marketing Theory and Practice* 3(2):64–74.

Wu D, Ray G, Geng X, Whinston A (2004) Implications of reduced search cost and free riding in e-commerce. *Marketing Science* 23(2):255–262.

Xia J, Niu W (2019) Adding clicks to bricks: An analysis of supplier encroachment under service spillovers. *Electronic Commerce Research and Applications* 37:100876.

Young J (2020) Us ecommerce sales grow 14.9% in 2019. https://www.digitalcommerce360.com/article/us-ecommerce-sales/, (Accessed on 06/24/2020).

Zhang J, Cao Q, He X (2020) Manufacturer encroachment with advertising. *Omega* 91:102013.

Zimmerman A (2012) Can retailers halt 'showrooming'. *The Wall Street Journal* 259(1):B1–B8.

INTEGRATING GLOBAL AND CLOSED-LOOP SUPPLY CHAINS TO MINIMIZE COSTS AND GHG EMISSIONS USING LPP

Hiromasa Ijuin, hju1n@yahoo.ne.jp; The University of Electro-Communications, Tokyo, Japan, +81-42-443-5269
Yuki Kinoshita, ykinoshita.rs.0401@gmail.com; The University of Electro-Communications, Tokyo, Japan, +81-42-443-5269
Tetsuo Yamada*, tyamada@uec.ac.jp; The University of Electro-Communications, Tokyo, Japan, +81-42-443-5269
Surendra M. Gupta, S.Gupta@northeastern.edu, Northeastern University, Boston, U.S.A, +1-617-373-4846
*Corresponding Author

ABSTRACT

The 27th session of the Conference of the Parties to the United Nations Framework Convention on Climate Change (COP 27) confirmed their commitment to limit global temperature rise to 1.5 degrees Celsius above pre-industrial levels. In order to decrease the GreenHousse Gas (GHG) emission, the manufacturing companies are required to procure components from suppliers with lower GHG emission and to recycle the End-of-Life (EOL) products for recovering GHG emission.

The global supply chain network transports the component and assembly product to factory or market. Based on Life Cycle Inventory (LCI) database, the amount of GHG emissions for procuring components is dependent on material production countries because the energy mix of power sources is different from countries. On the other hand, the reverse supply chain network transports and recycles EOL products. Reverse supply chain enables us to encourage the resource circulation and the GHG suppression. Thus, manufacturing companies have to integrate their Global Closed-Loop Supply Chain (GCLSC) network design environmentally-friendly and economically. However, environmentally-friendly and economical GCLSC sometimes are in conflict. Therefore, Decision Maker (DM) should apply multi-criteria decision making methods for designing GCLSC. Linear Physical Programming (LPP) is one of solving method for multi-criteria decision making.

This study designs the GCLSC network for material-based GHG emissions and costs using LPP. First, the GCLSC is modeled to determine each facility on that network. Second, the objective functions are formulated to minimize the GHG emissions and the total cost using LPP. Finally, numerical experiments are conducted in a case of vacuum cleaner.

Keywords: Closed-Loop Supply Chain, Global Supply Chain, Carbon Recovery, Multi-Criteria Decision Making, Linear Physical Programming

1. INTRODUCTION

In 2022, the 22th session of the Conference of the Parties to the United Nations Framework Convention on Climate Change (COP 27) [1] was held. In this previous conference, governments from all participating countries reaffirmed efforts to reduce GreenHouse Gas (GHG) emissions to control the global average temperature increase to under 1.5 Celsius degrees [1]. The GHG emissions of components and products in production differ among in each country (Urata et al. (2017)) [2]. To decrease GHG emissions in products, manufactures are required to procure parts from overseas suppliers with lower GHG emissions and recycle End-of-Life (EOL) products to recover the amount of embedded GHG emissions by recycling. Thus, manufacturing companies must establish their own global closed-loop supply chain (GCLSC) networks. The GCLSC includes a global forward supply chain (Urata et al. (2017)) [2] and reverse supply chain (Ijuin et al. (2019)) [3]. The global supply chain is a network for purchasing

components, assembling products, and transporting them to a demand city (Ravindran and Warsing Jr (2013)) [4]. The reverse supply chain is a network that collects EOL products from users and recycles them to recover materials (Battaïa and Gupta (2015)) [5]. Furthermore, recycling EOL products can reduce material-based GHG emissions because the use of recovered materials in the manufacture of products can reduce GHG emissions when new products are manufactured from new resources according to Life Cycle Assessment (LCA) (Itsubo et al. (2007)) [6]. The decision maker (DM) for designing a GCLSC network must have responsibility environmentally friendly and cost-effective. However, environmentally friendly and cost-effective GCLSCs are sometimes conflicting goals such as reducing material-based GHG emissions and costs, and Research Questions (RQs) are occur as follows:

RQ1. How to reduce and recover material-based GHG emissions decided by both of the procurement components and the recycling EOL product stages in designing GCLSC networks?

RQ2. Which components should be switched for conflicting goals such as GHG emissions and costs minimization.

Therefore, the DM should apply multi-criteria decision making methods when designing GCLSCs.

Linear Physical Programming (LPP) (Messac et al. (1996)) [7] is one of the methods for solving multi-criteria decision making problems, where DMs can reflect their preferences in ranges of each objective function, and remove the effort of weight parameters for each objective function.

Kazemia et al. (2019) [8] investigated the studies in the area of reverse logistics (RL) and closed-loop supply chain management (CLSCM). As gaps and research opportunities based on their content analysis of RL and CLSCM, they mentioned nine papers that addressed environmental concerns in RL and CLSCM. However, they pointed out that their studies have been to translate environmental aspects into costs. Polat et al. (2019) [9] design the integration of forward and reverse logistics network for an allocation and hub location selection problem about an electrolytic wire company by using mixed integer linear programming. Their model intended to reduce transportation costs, improve the efficiency of transportation activities, increase customer satisfaction, expand customer portfolio, and maintain the competitive advantage of the company. However, the objective function of their model is only cost. Therefore, their model analyzed only financial aspect. Dulman and Gupta (2018) [10] presented a model for closed-loop supply chain study that included maintenance operations and EOL product processing. However, they did not consider the evaluation of environmental aspect such as GHG emissions. Ijuin et al. (2021) [11] designs the GCLSC network to minimize GHG emissions and total cost on global and local supply chain network. However, they used ε constraint for GHG emissions. Thus, they did not use multi-criteria decision making methods.

This study models the GCLSC for material-based GHG emissions and costs using LPP. First, problem example of GHG emissions and procurement cost among nations and experimental scenarios are set for designing in the GCLSC network. Next, the results are shown and discussed in terms of the GHG emissions, the costs, and the supplier selection based on each scenario. Finally, objective functions are described and formulated to minimize GHG emissions and the total cost using LPP. The contribution of this paper is that the GHG reduction by changing supplier selection is more effective than one by recycling EOL products in the experiments.

2. MODEL &FOMULATION

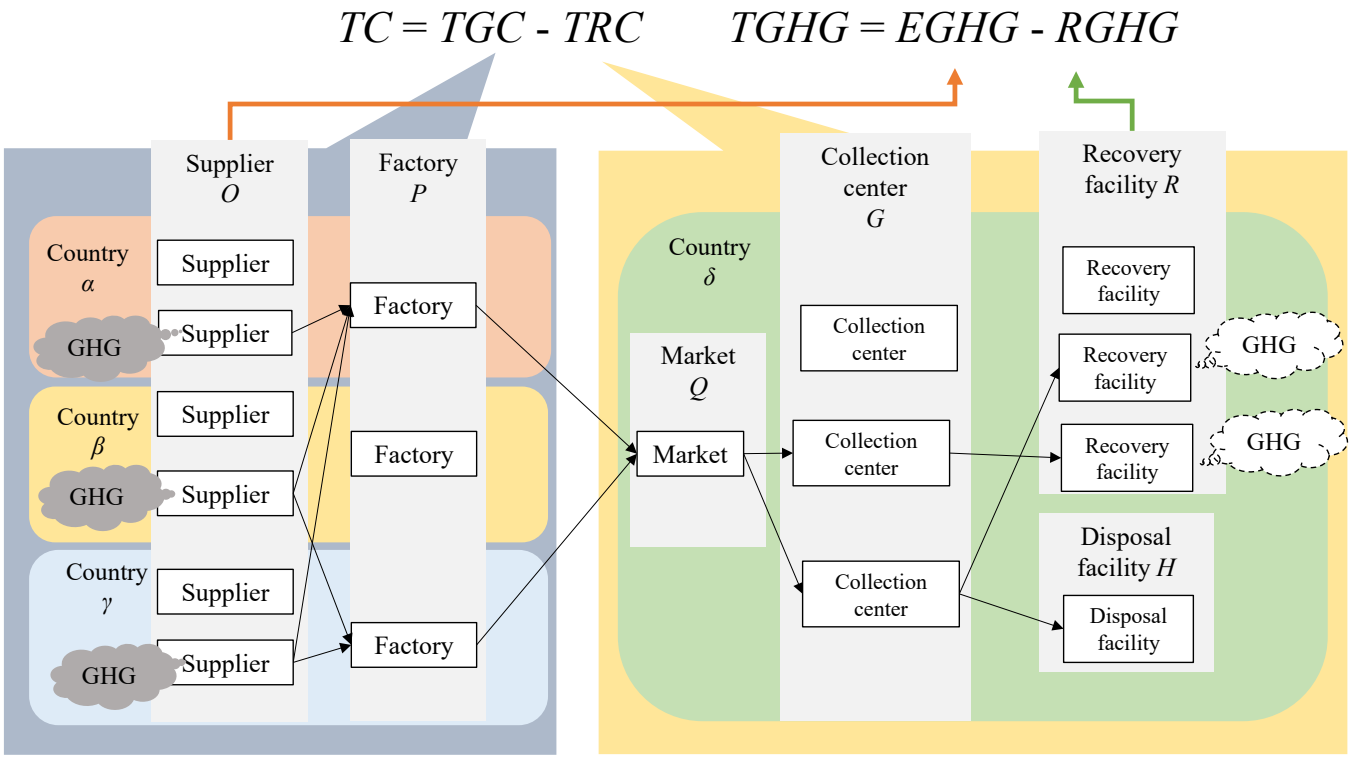


Figure 1: Model of the integrating global and closed-loop supply chains with EOL product status

Similar to Ijuin et al. (2021) [11], the GCLSC network prepared in this study integrates a global supply chain network (Urata et al. (2019)) [2] and a local reverse supply chain network (Ijuin et al. (2019)) [3]. The global supply chain network operates to procure and transport component j from supplier o, manufacture assembly products in factory p, transport component j from supplier o to factory p and assembled products from factory p to market q. In market p, assembly products are sold to the user and returned by the user as EOL products. The amount of material-based GHG emissions depends on the country in which the supplier l is located. Therefore, the GHG emissions GHG_k^{RS} of component j are generated based on supplier l in the global supply chain.

The local reverse supply chain collects EOL products from market q and transports them to the collection center g. The collection center g identifies the EOL products and divides each status k based on the user's situation. The status of the EOL product affects the recycling cost, C_k^{RC}, and recovered GHG emissions volumes GHG_k^{RS} by recycling EOL product. Subsequently, some EOL products with status k are transported to recovery facility r for recycling. Moreover, the remaining EOL products are transported and disposed in the disposal facility h. In this case, GHG emissions are not recovered and saved by disposed EOL products.

Similar to Ijuin et al. (2021) [11], this study sets two objective functions for minimizing material-based GHG emissions and the total cost using integer programming (Hilier and Lieberman (2005)) [12].

The variables and parameters of the mathematical formulation are listed as follows:

i) Sets

I : Set of objective functions

O : Set of suppliers

P : Set of factories

Q : Set of markets

G : Set of collection centers

R : Set of recycling facilities

H : Set of disposal facilities

J : Set of components

K : Set of statues

ii) Indices

i : Index of objective function ($i \in I$)

s : Index of desirability ranges (s=2, 3, 4, 5)

o : Index of supplier ($o \in O$)

p : Index of factory ($p \in P$)

q : Index of market ($q \in Q$)

g : Index of collection center ($g \in G$)

r : Index of recovery facility ($r \in R$)

h : Index of disposal facility ($h \in H$)

j : Index of components ($j \in J$)

k : Index of status for the EOL products ($k \in K$)

iii) Parameters

$\widetilde{w}_{is}^{+}$: Positive deviation weight of the s^{th} range of objective i

t_{is}^{+} : Positive limit to the s^{th} range of objective i

C_{oj}^{PC} : Procurement cost of components j from supplier o [US$]

C_{p}^{MC} : Manufacturing cost of an assembly product at factory p[US$]

C_{k}^{RC} : Recycling cost of EOL product with status k [US$]

C_{op}^{LC} : Transportation cost between nodes o, p [US$]

C_{pq}^{LC} : Transportation cost between nodes p, q [US$]

C_{qg}^{LC} : Transportation cost between nodes q, g [US$]

C_{gr}^{LC} : Transportation cost between nodes g, r [US$]

C_{gh}^{LC} : Transportation cost between nodes g, h [US$]

C_{p}^{FC} : Opening facility cost of node p [US$]

C_{r}^{FC} : Opening facility cost of node r [US$]

GHG_{oj}^{PE} : GHG emissions of component j procured from supplier o [g-CO_2eq]

GHG_{k}^{RE} : Recovered GHG emissions by recycling EOL product with status k [g-CO_2eq]

N_{q}^{DN} : Number of demands in market q[units]

N_j^{CN} : Required number of component *j* to produce an assembly product [units]

S_{oj} : 1, when supplier *o* can supply component *j*
0, otherwise

N_{rk}^{AN} : Number of available production capacity of EOL product with status *k* at recovery facility *r* [units]

N_p^{PN} : Total number of production capacity at node *p* [units]

N_r^{PN} : Total number of production capacity at node *r* [units]

iii) Variables

v_{opj} : Number of components *j* transported between nodes *o*, *p* [units]

v_{pq} : Number of products/EOL products transported between nodes *p*, *q* [units]

v_{qg} : Number of products/EOL products transported between nodes *q*, *g* [units]

v_{grk} : Number of EOL products transported between nodes *g*, *r* with status *k* [units]

v_{ghk} : Number of EOL products transported between nodes *g*, *h* with status *k* [units]

z_{oj} : Number of component *j* provided from supplier *o*

z_p : Number of products manufactured at factory *p* [units]

z_{gk} : Number of EOL products with status *k* collected at collection center g [units]

z_{rk} : Number of EOL products with status *k* transported at node *x* [units]

u_r : Binary value; 1 if node *x* is opened, else 0

$d_{i,s}^{+}$: Positive deviation variable from the s^{th} range limit of objective *i*

TC : Total cost of global closed-loop supply chain network [US$]

TGC : Total cost of global supply chain network [US$]

TRC : Total cost of local reverse supply chain network [US$]

TGHG : Total GHG emissions [g-CO_2eq]

EGHG : Total GHG emissions by procured components [g-CO_2eq]

RGHG : Total recovered GHG emissions weight by recycling EOL product [g-CO_2eq]

The total GHG emissions is the difference between procurement of component *j* from supplier *o* and recycling EOL products with status *k* in Equation (1). The total cost *TC* indicates the sum of the total cost of the global supply chain network and the closed-loop reverse supply chain network, as shown in Equation (2).
Objective function:

$$TGHG = EGHG - RGHG \quad (1)$$

$$TC = TGC + TRC \quad (2)$$

Equation (3) expresses the GHG emission by procured components. Total GHG emissions *EGHG* is calculated based on the per for GHG emissions GHG_{oj}^{PE} and the number of components z_{oj} provided from supplier *o*. On the other hand, the recovered GHG emissions defines Equation (4). Total recovered GHG emissions *RGHG* is multiped by the per for recovered GHG emissions GHG_k^{RE} with status *k* and the number of EOL products z_{rk} recycled at recovery facility *r*.

Equation (5) shows the total cost of the global supply chain *TGC*, including below costs: the procurement cost C_{oj}^{PC} of component j from supplier o, manufacturing cost C_p^{MC} of product in factory p, transporting cost C_{op}^{LC} for transportation between supplier o and factory p, transporting cost C_{pq}^{LC} for transportation between factory p and market q, opening cost C_p^{MC} of factory p. Equation (6) shows the total cost of the local reverse supply chain *RGC*, consisting several costs: the recycling cost C_k^{RC} of the EOL product with status k, the opening facility cost C_r^{FC} of recovery facility r, and the transportation costs C_{qg}^{LC}, C_{gr}^{LC}, C_{gh}^{LC} from market p to collection center g, from collection center g to recovery facility r, from collection center g to disposal facility h, respectively. Equation (7) expresses the constraint that the supplier o can provide only a certain kind of components S_{oj} based on their production ability and that the number of required components type j at each factory is satisfied from proper suppliers. Equation (8) means all manufactured products at each facility p are sent to markets. The demand in each market q must be satisfied as shown in Equation (9). The transported products from each factory p must be equal to or lower than its production capacity as shown in Equation (10). Equation (11) means that the number of transported products to market is equal to one of transported EOL product to collection center. Equation (12) shows that all EOL products collected from market are transported to collection center g. Equations (13) and (14) show that all EOL products in collection center g are transported to recovery facility r or disposal facility h. The transported EOL products from each recovery factory r must be equal to or lower than its production capacity N_{rk}^{AN} and N_r^{PN} based on EOL product status and total capacity as shown in Equations (15) and (16).

Constraints:

$$EGHG = \sum_{o\in O}\sum_{j\in J} GHG_{oj}^{\mathrm{PE}} z_{oj} \tag{3}$$

$$RGHG = \sum_{r\in R}\sum_{k\in K} GHG_{k}^{\mathrm{RE}} z_{rk} \tag{4}$$

$$TGC = \sum_{o\in O}\sum_{p\in P}\sum_{j\in J} C_{op}^{\mathrm{LC}} v_{opj} + \sum_{p\in P}\sum_{q\in Q} C_{pq}^{\mathrm{LC}} v_{pq} + \sum_{o\in O}\sum_{j\in J} C_{oj}^{\mathrm{PC}} z_{oj} + \sum_{p\in P} C_p^{\mathrm{MC}} z_p + \sum_{p\in P} C_p^{\mathrm{FC}} u_p \tag{5}$$

$$TRC = \sum_{q\in Q}\sum_{g\in G} C_{qg}^{\mathrm{LC}} v_{qg} + \sum_{g\in G}\sum_{r\in R}\sum_{k\in K} C_{gr}^{\mathrm{LC}} v_{grk} + \sum_{g\in G}\sum_{h\in H}\sum_{k\in K} C_{gh}^{\mathrm{LC}} v_{ghk} + \sum_{r\in R}\sum_{k\in K} C_{rk}^{\mathrm{RC}} z_{rk} + \sum_{r\in R}^{R} C_r^{\mathrm{FC}} u_r \tag{6}$$

$$\sum_{o\in O} S_{oj} v_{opj} = NP_j z_p \qquad \forall p \in P, \forall j \in J \tag{7}$$

$$\sum_{q\in Q} v_{pq} = z_p \qquad \forall p \in P \tag{8}$$

$$\sum_{p\in P} v_{pq} = N_q^{\mathrm{DN}} \qquad \forall q \in Q \tag{9}$$

$$\sum_{q\in Q} v_{pq} \le N_p^{\mathrm{PN}} u_p \qquad \forall p \in P \tag{10}$$

$$\sum_{g\in G} v_{qg} = N_q^{DN} \qquad \forall q \in Q \tag{11}$$

$$\sum_{q\in Q} v_{qg} = \sum_{k\in K} z_{gk} \qquad \forall g \in G \tag{12}$$

$$\sum_{g\in G} v_{grk} = z_{rk} \qquad \forall r \in R, \forall k \in K \tag{13}$$

$$\sum_{r\in R} v_{grk} + \sum_{h\in H} v_{ghk} = z_{gk} \qquad \forall g \in G, \forall k \in K \tag{14}$$

$$\sum_{g\in G} v_{grk} \le N_{rk}^{\mathrm{AN}} \qquad \forall r \in R, \forall k \in K \tag{15}$$

$$\sum_{g\in G}\sum_{k\in K} v_{grk} \le N_r^{\mathrm{PN}} u_r \qquad \forall r \in R \tag{16}$$

$$v_{opj}, v_{pq}, v_{qg}, v_{grk}, v_{ghk}, z_{oj}, z_p, z_{gk} \ge 0 \tag{17}$$

2.1 LPP objective functions

To harmonize the total GHG emissions and cost, this study applies LPP (Messac et al. (1996)) [7]. The aggregate objective function for GHG emissions and the total cost is shown in Equation (18). Objective function i is divided in five limits $t_{i,s}^{+}$(s = 2, 3, 4, 5) and six ranges called "Ideal" (Under $t_{i,1}^{+}$), "Desirable" ($t_{i,1}^{+} \sim t_{i,2}^{+}$), "Tolerable" ($t_{i,2}^{+}$

~ $t_{i,3}^{+}$), "Undesirable" ($t_{i,3}^{+}$ ~ $t_{i,4}^{+}$), "Highly-Undesirable" ($t_{i,4}^{+}$ ~ $t_{i,5}^{+}$), and "Unacceptable" (Over $t_{i,5}^{+}$). The weight parameter $\widetilde{w}_{is}^{+}$ indicates the weight value for objective function *i* between each range $t_{i,s}^{+}$ – $t_{i,s-1}^{+}$. The deviation d_{is}^{+} shows the difference in the value of objective function *i* and each limit $t_{i,s}^{+}$.

$$\sum_{i\in I}\sum_{s=\{2,3,4,5\}}\widetilde{w}_{is}^{+}d_{is}^{+}\rightarrow min \tag{18}$$

The first objective function of this study is related to the total GHG emissions *TGHG* of the entire GCLSC network. The total *GHG* emissions *TGHG* is formulated as the constraint equations (19) and (20). By minimizing the *s*th deviation variable $d_{1,s}^{+}$ in equation (19), the aim is for the total GHG emissions *TGHG* to reach the ideal range $t_{1,s}^{+}$. The total GHG emissions *TGHG* is lower than the limit of unacceptable $t_{1,5}^{+}$, as shown in equation (20).

$$TGHG - d_{1,s}^{+} \le t_{1,s-1}^{+}. \qquad s = 2,\dots,5 \tag{19}$$

$$TGHG \le t_{1,5}^{+}. \tag{20}$$

The second objective function of this study is related to the total cost *TC* of the entire GCLSC network. The total cost *TC* is formulated as the constraint equations (21) and (22). By minimizing the *s*th deviation variable $d_{1,s}^{+}$ in equation (21), the aim is for the total cost *TC* to reach the ideal range $t_{2,s}^{+}$. The total cost *TC* is lower than the limit of unacceptable $t_{2,5}^{+}$, as shown in equation (22).

$$TC - d_{2,s}^{+} \le t_{2,s-1}^{+}. \qquad s = 2,\dots,5 \tag{21}$$

$$TC \le t_{2,5}^{+}. \tag{22}$$

3. PROBLEM EXAMPLE

3.1 GCLSC network and product example

Similar to Ijuin et al. (2021) [11], this study sets a GCLSC example where 52 cities in the US, Malaysia, China, and Japan are selected as candidates of suppliers and factories.

Material-based GHG emissions and procurement cost for component is dependent on each country. Table 1 shows the average component's material-based GHG emissions and procurement cost among the US, Malaysia, China, and Japan (Hayashi et al. (2020)) [13]. From this table, the Malaysian or Chinese components have high GHG emissions but low procurement costs. American or Japanese component have low GHG emissions but high procurement cost.

Table 1. Comparison of the average component's procurement GHG emissions and cost between the US, Malaysia, China, and Japan

Nations	Average procurement cost for component [US $]	Average GHG emissions for component [g-CO_2eq]
The US	0.09	103.58
Malaysia	0.07	221.04
China	0.08	556.56
Japan	0.13	91.72

3.2 Scenario of experiments

To evaluate the GCLSC network in terms of GHG emissions and cost, three scenarios are set in Table 2 as follows:

Scenario 1 - Total cost minimum: This scenario minimizes the total cost *TC* on GCLSC.

Scenario 2 - GHG emissions minimum: This scenario minimizes total GHG emissions *TGHG* by procuring components and recycling EOL products on the GCLSC.

Scenario 3 - GHG emissions and cost minimum using LPP: This scenario harmonizes the total GHG emissions *TGHG* and the total cost *TC* using LPP on the GCLSC.

Table 2. Scenarios for designing GCLSC network

Scenarios	Scenario Name	Cost reduction	GHG reduction
Scenario 1	Total cost minimum	□	
Scenario 2	Total GHG emissions minimum		□
Scenario 3	Total GHG emissions and total cost minimum using LPP	□	□

4. RESULTS

4.1 Result of costs and GHG emissions

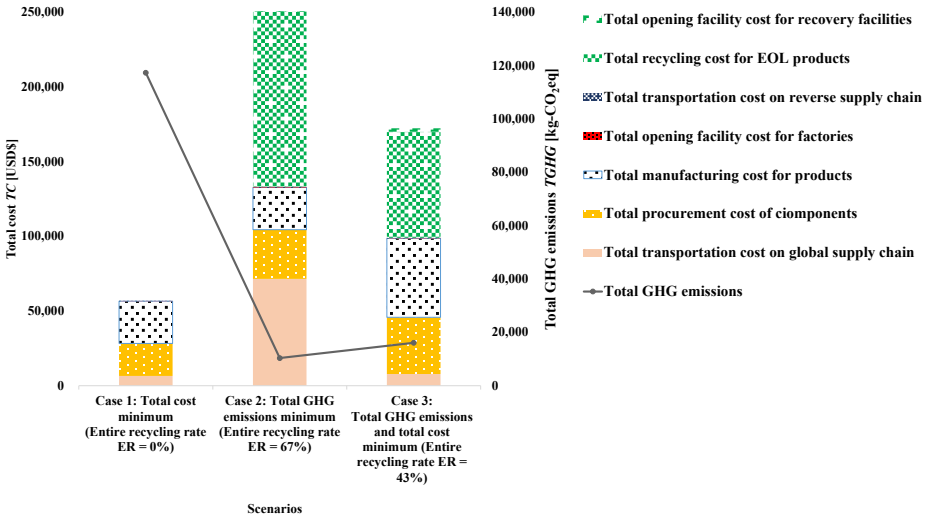


Figure 1 Result of global closed-loop supply chain (GCLSC) network in three scenarios

Figure 1 shows the results of the GCLSC network for the three scenarios. In scenario 1 (total cost minimum), the total cost *TC* is the lowest among the scenarios and is lower than ones of scenarios 2 (total GHG emissions minimum) and 3 (total GHG emissions and total cost minimum) by 228,715 [US$] (403% low) and 115,450 [US$] (203% low), respectively. However, the total GHG emissions were higher than ones in scenarios 2 and 3 by 106,896 [kg-CO2eq] (91% high) and 101,158 [kg-CO_2eq] (86% high), respectively. One reason for this is that components in case 1 are procured from China or Malaysia, which have the high GHG emissions as shown in Table 1. Furthermore, since this scenario does not recycle any EOL products, GHG emissions are not reduced by recycling the EOL products.

In scenario 2 (total GHG emissions minimum), total GHG emissions *TGHG* were lower than in scenario 3 by 5,738 [kg-CO_2eq] (55% low). However, the total cost *TC* was higher than scenario 3 by 113,264 [US$] (40% low). In this scenario, the components of the product were procured from Japan, the US, and Malaysia, which are countries with low GHG emissions.

On the other hand, scenario 3 obtained a harmonized solution by procuring components from countries with low GHG emissions and by disposing (not recycling) EOL products.

5. CONCLUSION AND FUTURE STUDY

This study modeled a global closed-loop supply chain network in an environmentally friendly and economically using LPP. The proposed model can contribute to the reduction of GHG emissions, thus fulfilling the key goal of

circular economy for GHG emissions and cost.

The novelty of this paper is that a GCLSC network is modeled where total costs can be reduced while reducing material-based GHG emissions embedded in products in both component procurement in forward supply chain and EOL product recycling in reverse supply chain. The contribution of this paper is to decide the numbers of components for procurement component, transportation, manufacturing on each route, and EOL product processed on each route, and facility locations for manufacturing and recycling by connecting global closed-loop supply chain networks. As the result, the changing supplier selection is more effective than increasing the number of recycled products.

Future research should consider other important objectives such as quality and delivery. Furthermore, a comparison with LPP and other solving methods for multi-criteria decision making should be conducted in the example.

ACKNOWLEDGEMENTS

This research was partially supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI, Grant-in-Aid for Scientific Research (A), JP18H03824, from 2018 to 2023 and Grant-in-Aid for JSPS Research Fellow, in 2023.

REFERENCE

[1] European Commission, “Global Energy and Climate Outlook 2021: Advancing towards climate neutrality”, https://publications.jrc.ec.europa.eu/repository/handle/JRC126767 (accessed on January 8, 2023)

[2] Urata T, Yamada T, Itubo N, Inoue M: Global supply chain network design and Asian analysis with material-based carbon emissions and tax. Computers and Industrial Engineering 2017, 113: 779-792

[3] Ijuin H, Kinoshita Y, Yamada T, Ishigaki A, Inoue M: Designing reverse supply chain network with costs and recycling rate by using linear physical programming. International Journal of Smart Computing and Artificial Intelligence 2019, 3(2): 57-76

[4] Ravindran A. R, Warsing Jr D. P: Supply Chain Engineering, CRC Press 2013

[5] Battaïa O, Gupta S. M.: Reverse supply chains: A source of opportunities and challenges. Journal of Manufacturing Systems 2015, 37: 587-588

[6] Itsubo N, Tahara K, Narita N: Introduction of LCA, Japan Environmental Management Association for Industry 2007 (in Japanese).

[7] Messac A, Gupta S. M, Akbulut B: Linear physical programming: A new approach to multiple objective optimization. Transactions on Operational Research 1996, 8: 39-59

[8] Kazemia N, Modakb N. M, Govindanc K: A review of reverse logistics and closed loop supply chain management studies published in IJPR: a bibliometric and content analysis. International Journal of Production Research 2019, 57(15-16): 4937-4960

[9] Polat O, Ünal Ö. F, Polat L. Ö, Güngör A: Responsible Manufacturing Issues Pertaining to Sustainability: Integrated Forward-Reverse logistics Network Design: An Application for Electrolytic Copper Conductor Reel Distribution, In: Responsible Manufacturing: Issues Pertaining to Sustainability. Edited by Alqahtani, A. Y, Kongar, E, Pochampally, K. K, Gupta S. M, Chapter 4, CRC Press, 2019, 85-103

[10] Dulman M. T, Gupta S. M: Evaluation of maintenance and EOL operation performance of sensor-embedded laptops. logistics 2022, 2(3): 22pages.

[11] Ijuin H, Yamada T: Problem of modeling global and closed-loop supply chain network design, 26th International Conference on Production Research. Taiwan (Online), 2021.

[12] Hillier F. S, Lieberman G. J: Introduction to operations research 8th edition, McGraw-Hill 2005

[13] Hayashi K, Matsumoto R, Yamada T, Nagasawa K, Kinoshita, Y: Designing a global supply chain network considering carbon tax and trans-pacific partnership agreement. The Society of Plant Engineering Japan Spring Conference 2020. Japan, 2020, 97–100. (In Japanese).

Simulation Analysis of Repair Kits with Customized Part Service Levels

John Maleyeff, maleyeff@bu.edu
Jia Fang, fangj@bu.edu
Jianbo Jiang, jjiang01@bu.edu

Boston University
Metropolitan College

Abstract

Public transportation entities like the MBTA in Boston maintain an inventory of parts needed to repair trains, buses, tracks, and a plethora of other equipment. Repairs are needed for equipment purchased over many decades resulting in a considerable inventory investment, although the demand for many parts is intermittent and procurement lead times can be long. This article describes the use of a repair kit simulation to explore the relationship between a part's service level and relevant inventory-related costs. It recognizes the unique aspect of many repairs because the work cannot commence unless all required parts are available. Parts vary in their costs and lead times, while repairs vary in their annual demand and part complexity. The analysis compares a constant service level for all parts in the repair kit to options that assign lower service levels for expensive parts or higher service levels for inexpensive parts. Results show these options to be favorable depending on the system's cost structure. The work is relevant to repair part managers at the MBTA and other systems that store millions of parts while adhering to tight budgets.

1. Introduction

The MBTA (Massachusetts Bay Transportation Authority), which operates the public transit system in Greater Boston, needs to maintain and repair buses, subways, ferries, tracks, stations, and other forms of infrastructure that may be decades old. This article concerns research that was motivated by challenges that exist when controlling the inventory for repair parts at the MBTA's central part warehouse. This warehouse stores over 10000 different parts that are used by a collection of repair garages located around the region. Almost all of the parts are subjected to intermittent demand because buses, subways and other capital equipment were purchased annually at relatively low volumes over many years. Lead times for their resupply are usually long, sometimes several months.

Determining the inventory policy for a part held in the warehouse is based on its lead time, along with the demand forecast for the repairs for which it is used. Repair part inventory managers need to set the target service level for each part. However, they are unaware of how these service levels should vary across parts that differ in cost, lead times, and demand rates. This research contributes to the development of inventory policies for a repair kit, defined as a set of parts needed to complete a specific repair job. The job cannot commence unless the entire repair kit is available.

In this article, a simulation is used to explore the relationship between part characteristics, their service levels, and relevant inventory costs. It is assumed that parts will only be sent to a garage when the entire collection of parts is available at the central warehouse. It follows prior work that

helped determine the list of assumptions for the inventory optimization model (Maleyeff et al, 2002). First, a k-means cluster analysis concluded that 80-90% of the repair parts had demand consistent with the Poisson distribution. Second, an analytical model was created to optimize parts but not entire repairs; this model suggests that part service levels should differ based on their lead times and part costs.

The article is organized as follows. A literature search is presented that frames the repair kit problem in the context of work published over the past decades of inventory modelling research. Then, a representative repair kit is presented whose flexible structure allows for changing key parameters to explore their impacts. The simulation model description follows, using the STRESS structure. Results are presented and discussed. The focus is on a comparison of constant part service levels versus customized service levels with more expensive parts assigned lower service levels and/or less expensive parts assigned higher service levels. The article concludes with suggestions concerning future work for repair kit inventory modeling researchers.

2. Literature Search

Brumelle and Granot (1993) describe the repair kit inventory problem in the context of a technician who travels to the repair's location. The technician takes the parts they anticipate needing and cannot complete the repair unless they have every required part. Various authors have modeled the repair kit problem under a variety of assumptions. Mamer and Smith (1985) assumed an inventory policy where parts are ordered one at a time, and the main costs are carrying costs and a penalty cost when a part is unavailable. Others have considered cost minimization with service level goals (Hu et al, 2018), or cost minimization based on carrying costs and poor service penalties (Teunter, 2006). Prak et al (2017) used a heuristics approach to repair kit inventory control due to the computational complexity of their models, while considering shortage cost minimization. Nie and Sheng (2009) addressed the minimization of backorders in a two-echelon repair system.

Demand for repairs occurs when an entity fails; this demand is often intermittent (Engelmeyer, 2015). The normal distribution rarely applies because the time between failures will often follow a Weibull or other right-skewed distribution (Schuh et al, 2015). Components in constant failure stages will have demand that follows a Poisson distribution (Hollier, 1980; Arts et al, 2016). At times, the Poisson model may have an unstable mean when: (a) components approach the end of their useful life (Slay and Sherbrooke, 1988), (b) external factors contribute to failure likelihoods (Jiang et al, 2020), or (c) stored items are perishable (Taleizadeh et al, 2013; Meisheri et al, 2022).

Although their usage volume may be low, setting the target service level for a repair part's inventory model has significant cost implications (Guide and Srivastava, 1997). These costs can be especially high in a system like the MBTA that stores millions of parts across over 10000 part numbers. Inventory modeling researchers have considered the determination of target service levels. They have been optimized based on ABC part classifications (Teunter et al, 2009), the part's contribution to product profit margins (Millstein et al, 2014), and a multi-criteria part criticality classification based in part on failure probabilities, replenishment times, and number of suppliers (Molenaers et al, 2012).

When parts are classified according to traditional ABC classification, A items are sometimes assigned the highest service levels (Hu et al, 2018). However, C items are also handled in ways

to ensure their availability because they do not contribute to high carrying costs (Basten and van Houtum, 2014). Moore (1996) suggested a service level differentiation based on: (a) low volume, critical; (b) high volume, high use; and (c) low volume, noncritical. Nicholson et al (2004) recommended different service levels for various forms of inventory and distribution systems. Care must be taken, however, to adjust the optimal service levels as circumstances change (Lolli et al, 2019).

3. Methodology

Each repair kit can be described by a bill-of-material (BOM) that lists the parts required to complete the repair and the purchase information about each part. At the MBTA, a few parts are very expensive while many parts are much less costly. For example, repairing a bus seat requires expensive fabric in conjunction with less expensive fasteners. Repairing a subway window requires expensive glass along with inexpensive caulking material. In public transportation systems, a part type is often dedicated to one specific repair (e.g., because small numbers of buses and other equipment are purchased every year). Therefore, most of the parts on the BOM are not common across many repair types and therefore part substitution is limited.

Consider the repair kit BOM shown in Table 1 (where n is the number of different parts in the repair kit). The total part cost is \$400, and the part lead times range from 4 weeks to 32 weeks. Part D represents an expensive part whose cost (C_D) will be assumed to constitute the majority of the \$400 total part cost for one repair. The lead time for Part D is assumed to be L_D weeks. The other parts will have equivalent costs so that the total cost of all parts needed for the repair is \$400. Because more than one unit of a specific type may be required, the "Part Cost" column represents the total cost for this part type. This structure is assumed so that the impact of the Part D lead time, Part D cost, and the total number of parts in the repair kit can be analyzed without loss of generality.

Table 1: Example 16-Part Repair Kit Bill-of-Material

Part	Lead Time (Weeks)	Part Cost
A	4	$(400-C_D)/(n-1)$
B	6	$(400-C_D)/(n-1)$
C	8	$(400-C_D)/(n-1)$
D	L_D	C_D
E	10	$(400-C_D)/(n-1)$
F	12	$(400-C_D)/(n-1)$
G	14	$(400-C_D)/(n-1)$
H	16	$(400-C_D)/(n-1)$
I	18	$(400-C_D)/(n-1)$
J	20	$(400-C_D)/(n-1)$
K	22	$(400-C_D)/(n-1)$
L	24	$(400-C_D)/(n-1)$
M	26	$(400-C_D)/(n-1)$
N	28	$(400-C_D)/(n-1)$
O	30	$(400-C_D)/(n-1)$
P	32	$(400-C_D)/(n-1)$

4. Repair Kit Simulation

The STRESS framework (Monks et al., 2019) is used to describe the simulation that was created to represent the inventory system for one repair kit. This framework imposes a standard template consisting of five parts (objective, logic, data, experimentation, and implementation & code access). It is designed to enhance the description of a simulation model and its usefulness.

4.1 Objective

The purpose of the repair kit simulation is two-fold. First, the simulation is designed for use by decision makers who desire to test the appropriateness of an inventory policy that sets service levels for each part in a repair kit (using the same service level for each part or using service levels that vary by part). Second, the simulation is designed for use by researchers to test part-based analytical models (because computational complexities preclude the development of a practical analytical model for a multi-period, multi-part repair kit).

4.2 Logic

Figure 1 describes the logical flow of the simulation, where the user enters the following information:

1. Annual repair demand rate
2. Lead time for Part D (the expensive part)
3. Percentage of Part D's cost relative to the total $400 repair kit part cost
4. Number of different parts required for the repair
5. Service levels for each part
6. Fixed ordering cost
7. Annual carrying cost percentage

The simulation assumes a periodic review (i.e., weekly) inventory policy, where s_j is the reorder point for part j and S_j is the order-up-to level for part j. Reorder points are based on the part's service level and the distribution of its demand during the lead time, while the order-up-to level is based on the fixed ordering cost and the annual carrying cost percentage. The simulation assumes Poisson demand for repairs, and backlogged repairs are done when parts become available (at no extra cost). The simulation is run for a specified number of trials (i.e., macro replicates) and iterations (i.e., weeks) for each trial. Finally, a warmup period (number of weeks) is specified. The simulation generates the following outcomes:

1. The number of repair kits demanded and completed, and their ratio.
2. The estimated and 95% confidence interval for the mean annual inventory carrying cost.
3. The estimated and 95% confidence interval for the mean annual ordering cost.
4. The estimated and 95% confidence interval for the mean repair delay (number of weeks from repair demanded to repair completed).
5. The estimated and 95% confidence intervals for the overall repair service level (likelihood that the repair is completed in the week demanded).
6. The estimated and 95% confidence intervals for the overall repair fill rate (percentage of repairs done in the week they are demanded).

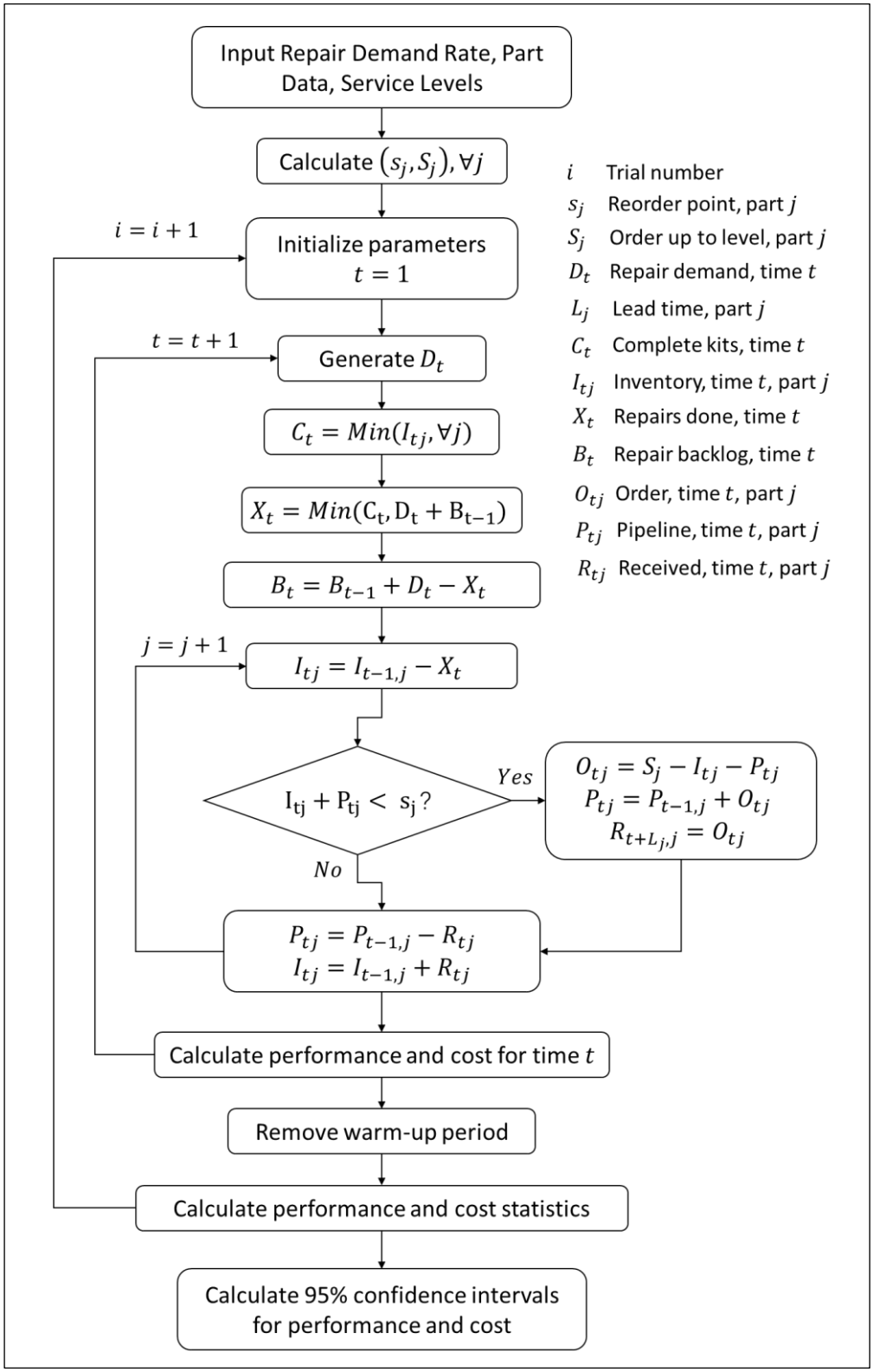


Figure 1: Simulation Flowchart

4.3 Data

The rationale that formed the basis of the simulation's assumptions and ranges of potential parameter values was based on prior analyses of data provided by the MBTA. These data justified the assumption of Poisson demand at intermittent levels (i.e., low rates), the ranges of procurement lead times (which ranged from a few weeks to several months), and the part cost ranges. The repair kit BOM's were created based on a range of parts per kit for the types of repairs undertaken at the MBTA.

4.4 Experimentation

The simulation is non-terminating. For the results detailed below, a warm-up of 300 iterations (i.e., weeks) ensured that the system was stabilized. This period was determined using Welch's method (Welch, 1983, pp. 268-328). The simulation was run for 100 trials, with 3000 iterations (i.e., weeks) per trial. The number of trials was chosen so that the standard error of the annual cost was approximately 1% of the average output values.

4.5 Implementation & Code Access

The simulation was programmed on a laptop with Windows 10 Enterprise, Version 21H1, OS Build 19043.2130. The simulation was developed in Python 3.9.12, a programming language that is offered free of charge by the Python Software Foundation. It incorporates packages numpy, pandas, scipy.stats, and matplotlib. Processing time per scenario on a Lenovo ThinkPad Carbon X1 laptop (with an Intel Core i7-1165G7 processor) took 5.36 minutes (16-part repair kit) and 8.66 minutes (32-part repair kit).

5. Results

The simulation was used to explore the performance of various inventory policies for the example repair kit (Table 1). The analysis focused on comparing the use of a constant service level for each part in the repair kit to alternatives that customized part service levels for parts in the same repair kit. Eight scenarios were used to run the half fraction factorial experimental design, shown in Table 2. The lead time for Part D (L_D) was assumed to be either 10 weeks or 30 weeks, with its cost assumed to comprise 60% or 85% of the total part cost for one repair. The repair demand rate was assumed to be 10 or 30 repairs per year, and the total number of parts in the repair kit was assumed to be 16 or 32.

Table 2: Experimental Design

Scenario	Annual Demand	Part D Lead Time	Part D Percent	# Parts in Repair
1	10	10	60	16
2	30	30	60	16
3	30	10	85	16
4	10	30	85	16
5	30	10	60	32
6	10	30	60	32
7	10	10	85	32
8	30	30	85	32

For this preliminary analysis, the controlled experiment was designed to explore the impact of the four key parameters. The objective was to evaluate the appropriateness of applying a customized set of service levels based on each part's unique characteristics and comparing the customized approach to a naïve approach that applies the same service level for each part. The simulation assumed a 24% annual inventory carrying cost, the fixed ordering cost was assumed to be zero (a reasonable assumption when orders are generated routinely and a large department processes many other transactions), and the cost of delay was assumed to be zero (a reasonable assumption when a buffer capacity of equipment exists).

5.1 Constant Service Level Analysis

The simulation was run for each of the 8 scenarios (Table 2), with part service levels held the same for each part in the repair kit. The service levels were incremented by 0.01 to document results for a range of potential inventory policies. The 95% confidence intervals for total carrying cost and the repair delay were computed and compared. Figure 2 is an example how the repair delays were affected by increases in part service levels for scenario 6 (where the annual demand was 10 repairs per year, the lead time for Part D was 30 weeks, Part D constituted 60% of the total part cost, and 32 parts were included in the repair kit). As expected, increases in service levels decrease repair delays, in this case monotonically as the service levels increase.

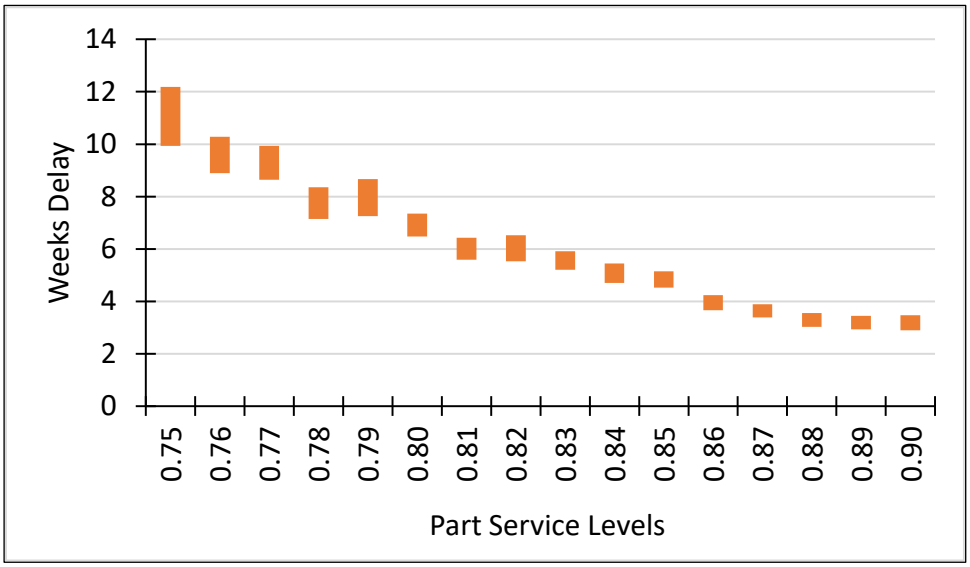


Figure 2: Repair Delay Confidence Intervals for Constant Part Service Levels (Scenario 6)

Figure 3 shows the 95% confidence intervals for annual carrying cost corresponding to scenario 6. As expected, the annual cost increases as service level increases. There is a distinct stepwise increase, however, when the part service levels increase from 0.82 and 0.83. This increase occurred because the order-up-to level for the high part cost item (Part D) increased by one (to 9 from 8) when the part service levels increased from 0.82 to 0.83, causing the disproportionate increase due to the very expensive Part D.

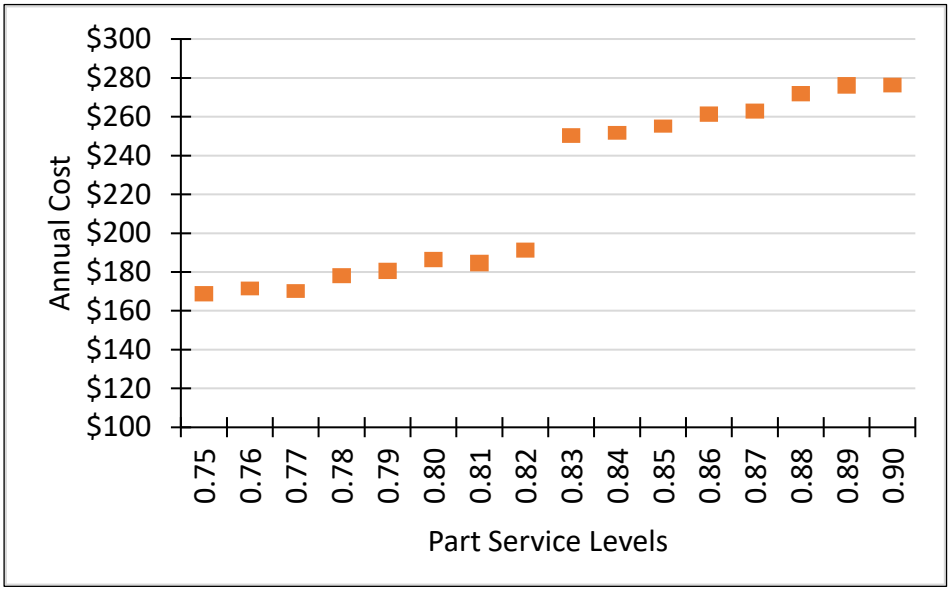


Figure 3: Annual Cost Confidence Intervals for Constant Part Service Levels (Scenario 6)

Because the cost of a delay is assumed to be zero, it is difficult to determine the best service level for parts in a repair kit. However, a maximum repair delay can be specified and the lowest alternative to achieve that specification can be found. Here, a 4-week maximum lead time was specified (modifications to this target delay will not affect the basic results reported below). The lowest set of part service levels whose upper confidence limit for repair delay was less than 4 weeks are shown in Table 3. It is noted that scenarios with higher annual demand tend to have lower service levels, regardless of the other three parameters.

Table 3: Summary of Results: Constant Service Levels

Scenario	Annual Demand	Part D Lead Time	Part D Percent	# Parts in Repair	Service Level	Weeks Delay 95% CI	Annual Cost 95% CI
1	10	10	60	16	0.87	2.9, 3.4	204, 209
2	30	30	60	16	0.79	3.4, 3.9	347, 362
3	30	10	85	16	0.78	3.3, 3.7	286, 292
4	10	30	85	16	0.87	2.8, 3.2	277, 285
5	30	10	60	32	0.80	3.2, 3.6	304, 313
6	10	30	60	32	0.87	3.4, 3.9	259, 267
7	10	10	85	32	0.87	3.3, 3.8	196, 199
8	30	30	85	32	0.79	3.5, 4.0	369, 383

5.2 Customized Service Level Analysis

Three alternative options were considered that adjusted service levels by either decreasing the service level of the expensive part (Part D), increasing the service level of the inexpensive parts, or both. Repair delays and annual carrying cost were compared for these alternatives. The following notation was used when describing the analysis results:

[0, 0] Keep the best constant service levels for entire repair kit (based on values in Table 3).

[0, -1] Keep the best constant service level for inexpensive parts and decrease service level for the expensive part (i.e., the order-up-to level for Part D decreases by 1).

[1, 0] Keep the best constant service level for the expensive part and increase service level for the inexpensive parts (i.e., the order-up-to levels for other parts increase by 1).

[1, -1] Increase the service level for inexpensive parts and decrease service level for the expensive part.

Figure 4 compares the performance and cost of the four alternatives for the combination of parameters corresponding to scenario 6 in Table 3. The [0,0] baseline (i.e., constant levels service for all parts) shows an estimated annual cost of about $263 and an estimated repair delay of about 3.6 weeks. For the alternative [0,-1], decreasing the order up-to-level by 1 for the expensive part decreases expected annual carrying cost by about 19.8% (to $211 from $263), while increasing the expected delay by about 1.2 weeks (to 4.8 weeks from 3.6 weeks). For the alternative [1,0], increasing the order-up-to level by 1 for the inexpensive parts increases expected annual carrying cost by about 14.8% (to $302 from $263), while decreasing the expected delay by about 1.6 weeks (to 2.0 weeks from 3.6 weeks). And for the alternative [1,-1], decreasing the order up-to-level by 1 for the expensive part and increasing the order-up-to level for inexpensive parts by 1 decreases expected annual carrying cost by about 6.1% (to $247 from $263), while increasing the expected delay by about 0.6 weeks (to 4.2 weeks from 3.6 weeks).

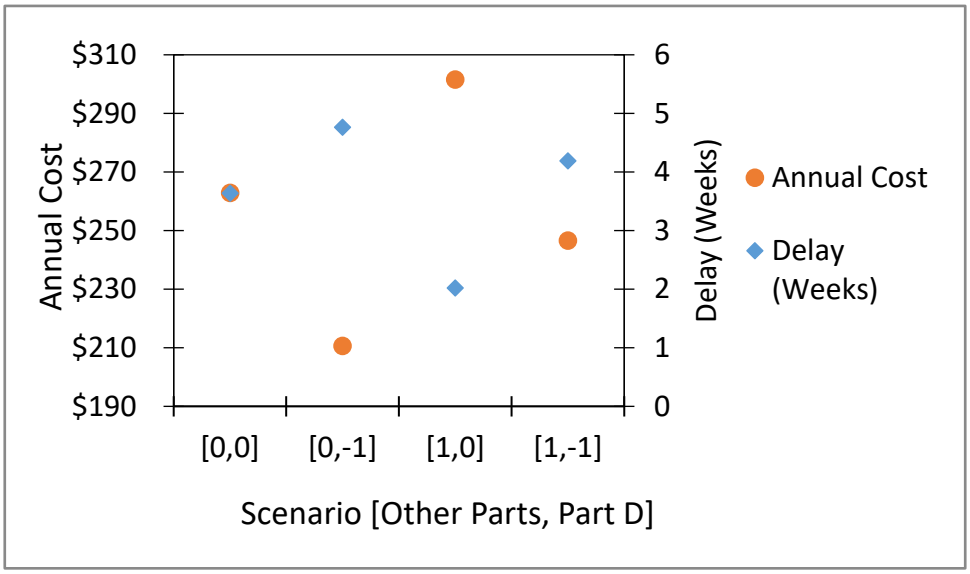


Figure 4: Comparison of Alternatives (Scenario 6)

Table 4 provides all of the results that compare the four alternatives, which are consistent across the eight scenarios. The preferable solution depends on the cost of delaying the repair. When repair delays are expensive (or otherwise undesirable), the best solution tends to be [1,0] because this option minimizes repair delays. When repair delays are inexpensive (or generally acceptable), the best solution tends to be [0,-1] because this option minimizes annual carrying costs. It is noted that both of these alternatives consist of applying a higher service level for inexpensive parts and a lower service level for expensive parts.

It is noted that cost differentials are more pronounced when the expensive part's service level is decreased for repairs that constitute a higher percentage of the total part cost in the repair kit. Repair delays decreases are greater when the expensive part's service level is decreased for scenarios where this part has a short lead time, regardless of annual repair demand or other parameters. The number of parts in a repair kit does not appear to impact preferred service levels; each scenario shows advantages of customizing part service levels.

Table 4: Summary of Results: Customized Service Levels

Scenario	Annual Demand	Part D Lead Time	Part D Percent	# Parts in Repair	Service Level	Alter-native	Weeks Delay	Annual Cost
1	10	10	0.60	16	0.87	[0, 0]	3.2	$206
						[0, -1]	4.2	$150
						[1, 0]	1.3	$247
						[1, -1]	3.9	$188
2	30	30	0.60	16	0.79	[0, 0]	3.6	$354
						[0, -1]	4.5	$294
						[1, 0]	2.3	$392
						[1, -1]	3.8	$334
3	30	10	0.85	16	0.78	[0, 0]	3.5	$289
						[0, -1]	3.6	$208
						[1, 0]	2.1	$303
						[1, -1]	2.4	$222
4	10	30	0.85	16	0.87	[0, 0]	3.0	$281
						[0, -1]	5.0	$195
						[1, 0]	2.0	$293
						[1, -1]	4.3	$207
5	30	10	0.60	32	0.80	[0, 0]	3.4	$309
						[0, -1]	3.6	$247
						[1, 0]	2.1	$341
						[1, -1]	2.4	$283
6	10	30	0.60	32	0.87	[0, 0]	3.6	$263
						[0, -1]	4.8	$211
						[1, 0]	2.0	$302
						[1, -1]	4.2	$247
7	10	10	0.85	32	0.87	[0, 0]	3.5	$197
						[0, -1]	4.3	$117
						[1, 0]	1.6	$213
						[1, -1]	3.8	$132
8	30	30	0.85	32	0.79	[0, 0]	3.8	$376
						[0, -1]	4.7	$292
						[1, 0]	2.7	$382
						[1, -1]	3.8	$308

6. Conclusion

Many inventory managers create inventory rules that assign higher service levels to critical parts, because of their disproportionate impact on the firm's success. It may appear wise for a repair part inventory manager to assign a higher service level to a critical part used in an equipment repair. The simulation employed in this article found that performance may be enhanced by assigning a lower service level to parts deemed critical because they constitute a high proportion of a total repair kit's part cost. The best service level for a part also appears to depend on other factors, such as the part's lead time and the annual repair demand. These results are important for a public transportation system like the MBTA in Boston where the total number of stored parts numbers in the millions and budget gaps are common.

The research reported here is ongoing. The simulation model currently assumes Poisson repair demand, and part lead times and other parameters are consistent with circumstances that exist at the MBTA. Some limitations will be addressed as the work continues. First, demand assumptions will be made more robust, because a constant failure rate would not apply for older equipment and the Poisson model does not apply to parts used for scheduled maintenance. Second, additional cost elements will be incorporated, such as fixed ordering costs and repair delay costs. Third, the impacts of substitutions and expediting on optimal service levels will also be studied.

7. References

Arts, J., Basten, R., & Van Houtum, G. (2016). Repairable stocking and expediting in a fluctuating demand environment: Optimal policy and heuristics. *Operations Research, 64*(6), 1285-1301.

Basten, R.J.I. & van Houtum, G.J. (2014). System-oriented inventory models for spare parts. *Surveys in Operations Research and Management Science, 19*, 34-55.

Brumelle, S. & Granot, D. (1993). The repair kit problem revisited. *Operations Research, 41*(5), 994-1006.

Engelmeyer, T. (2015). *Managing Intermittent Demand.* Wiesbaden: Springer Gabler.

Guide Jr., V.D.R. & Srivastava, R. (1997). Repairable inventory theory: Models and applications. *European Journal of Operational Research, 102*(1), 1-20.

Hollier, R.H. (1980). The distribution of spare parts. *International Journal of Production Research, 18*(6), 665-675.

Hu, Q., Boylan, J.E., Chen, H., & Labib, A. (2018). OR in spare parts management: A review. *European Journal of Operational Research, 266*(2), 395-414.

Jiang, A., Tam, K.L., Guo, X., & Zhang, Y. (2020). A new approach to forecasting intermittent demand based on the mixed zero-truncated Poisson model. *Journal of Forecasting, 39(1),* 69-83.

Lolli, F., Balugani, E., Ishizaka, A., Gamberini, R., Rimini B., & Regattieri A. (2019). Machine learning for multi-criteria inventory classification applied to intermittent demand, *Production Planning & Control, 30*(1), 76-89.

Maleyeff, J., Xu, J, Kaffaty, V., & Yang Y. (2022). Optimal service level for repair part inventory considering the cost of holding other parts. Proceedings of the 53rd Annual Conference of the Decision Sciences Institute, 722-734.

Mamer, J.W. & Smith, S.A. (1985). Job completion-based inventory systems: Optimal policies for repair kits and spare machines. *Management Science, 31*(6), 703-718.

Meisheri, H., Sultana, N.N., Baranwal, M. (2021). Scalable multi-product inventory control with lead time constraints using reinforcement learning. *Neural Computing & Applications, 34*(3), 173-175.

Millstein, M.A., Yang, L., & Li, H. (2014). Optimizing ABC inventory grouping decisions. *International Journal of Production Economics, 148*, 71-80.

Molenaers, A.N., Baets, H., Pintelon, L., & Waeyenbergh, G. (2012). Criticality classification of spare parts: A case study. *International Journal of Production Economics, 140*(2), 570-578.

Monks, T., Currie, C.S., Onggo, B.S., Robinson, S., Kunc, M., & Taylor, S.J. (2019). Strengthening the reporting of empirical simulation studies: Introducing the stress guidelines. *Journal of Simulation, 13*(1), 55-67.

Moore, R. (1996). Establishing an inventory management program. *Plant Engineering, 50*(3), 113+.

Nicholson, L., Vakharia, A.J., Selcuk Erenguc, S. (2004). Outsourcing inventory management decisions in healthcare: Models and application. *European Journal of Operational Research, 154*(1), 271-290.

Nie, T., & Sheng, W. (2009). Simulation Analysis of multi-echelon inventory for repairable items. *International Conference on Information Engineering and Computer Science*, 1-4.

Prak, D., Saccani, N., Syntetos, A., Teunter, R., & Visintin, F. (2017). The repair kit problem with positive replenishment lead times and fixed ordering costs. *European Journal of Operational Research, 261*(3), 893-902.

Schuh, P., Schneider, D., Funke, L., & Tracht, K. (2015). Cost-optimal spare parts inventory planning for wind energy systems. *Logistics Research, 8*(1), 1-8.

Slay, F.M., & Sherbrooke, C.C. (1988). The nature of the aircraft component failure process: A working note. Defense Technical Information Center, Virginia.

Taleizadeh, A.A., Mohammadi, B., Cárdenas-Barrón, L.E., & Samimi, H. (2013). An EOQ model for perishable product with special sale and shortage. *The International Journal of Production Economics, 145*(1), 318-338.

Teunter, R.H. (2006). The multiple-job repair kit problem. *European Journal of Operational Research, 175*(2), 1103-1116.

Teunter, R.H., Babai, M.Z., & Syntetos, A.A. (2009). ABC classification: Service levels and inventory costs. *Production and Operations Management, 19*(3), 343-352.

Welch, P.D. (1983). The statistical analysis of simulation results. In S. S. Lavenberg (Ed.), The computer performance modeling handbook. New York: Academic Press.

Top Management Tournament Incentives and Inventory Efficiency

Abstract: This paper investigates whether the level of tournament incentives for top executives is related to firm inventory efficiency using a sample of U.S. manufacturing companies. We find that greater pay dispersion among top management (our proxy for tournament incentives) is associated with higher efficiency in inventory management, measured by inventory levels, inventory turnover, and leanness. Our results also show that pay dispersion has a stronger effect on inventory efficiency when companies are subject to a greater degree of the bullwhip effect. However, the above evidence is only found in the pre-COVID-19 pandemic period (i.e., before March 2020). During the pandemic, the positive effect of pay dispersion on inventory efficiency disappears as companies experience greater operational uncertainty.

Key words: Tournament incentives, pay dispersion, inventory efficiency, bullwhip effect, COVID-19 pandemic

1. Introduction

This paper investigates the relationship between the tournament incentives for top executives and firm's inventory efficiency in manufacturing companies. Prior literature in accounting and finance show that stronger tournament incentives for top executives have an impact on various managerial decisions, resulting in higher profitability, better firm performance in the stock market, yet riskier firm policies, and thus higher stock return volatility and lower credit ratings (e.g., Kale et al., 2009; Kini and Williams, 2012; Bannister et al., 2020). Furthermore, a stream of operations management studies shows that reducing inventory holdings increases operational efficiency, thereby improving firm financial performance (e.g., Eroglu and Hofer, 2011; Barker et al., 2022). Therefore, managers motivated by tournament incentives are more likely to maintain higher inventory efficiency. However, there is a dearth of evidence on whether tournament incentives result in more efficient inventory management, one of the most critical operating policies. This study attempts to fill this void by examining the relation between top executives' tournament incentives and inventory efficiency. We first establish the link between top management tournament incentives and firms' inventory efficiency; then explore how such association varies depending on the bullwhip effect as well as the impact of COVID-19 pandemic.

Inventory management is a key factor in improving firm operational and financial performance (Chuang et al, 2019; Barker et al, 2022). Previous studies have shown that reducing inventory holding minimizes waste and thereby increases efficiency (Silver, 1981; Bonney, 1994; Womack et al., 2007). Although academics have extensively investigated inventory efficiency, which plays a critical role in the success of different types of firms (e.g., Eroglu and Hofer, 2011; Mishra, et al., 2013; Shan and Zhu, 2014; Lin et al., 2018; Almutairi et al., 2022), it remains unclear

what are the determinants of inventory efficiency. This is important as both academics and practitioners are interested in factors resulting in higher level of inventory efficiency. Accordingly, our study addresses this issue from the perspective of top management tournament incentives and investigates the association between tournament incentives and inventory efficiency.

Tournament theory suggests that workers are provided with promotion-based incentives within firms ((Lazear and Rosen 1981; Green and Stokey 1983). In order to induce effort, firms tend to increase the pay gap up the hierarchy and only the best relative performer is promoted to the next higher level in the organization. Such a tournament serves as the incentive for top executives to work harder and achieve better performance to increase their chances of promotion. In addition, tournament theory implies that promotion-based incentive motivates top managers to take riskier actions such as bearing more operational risk resulting from lower inventory holdings. As lower inventory levels benefit firms by reducing cost and therefore improving operational performance and financial performance, both performance-increasing incentive and risk-taking incentive could incentivize top executives to lower inventory levels. However, whether tournament incentives can improve inventory efficiency and thus firm performance is an empirical question.

Previous studies identify that bullwhip effect has a significant impact on inventory efficiency (e.g., Lee et al., 1997; Chandra and Grabis, 2005; Shan and Zhu, 2014; Wang and Disney, 2016; Osadchiy et al., 2021). Higher inventory level resulting from an enlarged bullwhip effect harms inventory efficiency and firms' operational performance. Motivated by larger tournament incentives, top executives might spend more effort and respond more proactively in inventory management to mitigate the negative impact of the bullwhip effect on inventory efficiency. Therefore, we hypothesize that the effect of top executive pay dispersion on inventory

efficiency is stronger for firms with a greater degree of bullwhip effect. In addition, starting from March 2020, the business world has experienced unprecedent uncertainty due to the COVID-19 pandemic. The extreme situation of the pandemic may prompt top managers to rethink inventory management and shift the focus from inventory efficiency to other more critical issues related to supply chain disruptions. Accordingly, we also explore whether the effect of top executive pay dispersion on inventory efficiency will be weakened during the COVID-19 pandemic.

The data in this study comes from SEC's EDGAR database of manufacturing firms in the United States. We focus on manufacturing firms as inventories account for a sizable portion of asset base for such firms. The sample consists of a total of 12,669 firm-year observations from 1,153 different manufacturing firms. Our results from multiple regression models support all the hypotheses. The empirical findings show a positive effect of top executive pay dispersion on inventory efficiency. Furthermore, this positive effect is stronger when firms are subject to a greater degree of the bullwhip effect. Finally, during the COVID-19 pandemic, our evidence shows a less pronounced relationship between top executive pay dispersion and inventory efficiency regardless of an increasing degree of the bullwhip effect.

To our best knowledge, this paper is the first empirical study that investigates inventory efficiency from the angle of top management tournament incentives. By establishing the link between the tournament incentives and inventory efficiency, this study not only adds to the research on inventory efficiency but also helps further understanding of managerial motivation driven by tournament incentives. Specifically, inventory efficiency, as an important indicator of operational performance, is of interest to all types of firms and their top executives. Along this line, our study identifies top management tournament incentives as a factor that is positively related to inventory efficiency. The findings contribute to the research on inventory management

and how to increase inventory efficiency. In addition, our evidence contributes to the research on tournament incentives by shedding light on specific operational decisions that tournament incentives motivate top managers to make in order to increase their chances of being promoted. Considering the bullwhip effect as well as the extreme situation of COVID-19 pandemic, this study enriches the research on top management tournament incentives from the perspective of inventory management.

Finally, our findings have practical implications and potentially provide guidance for practitioners. Specifically, we show that in stable business environment (i.e., outside the pandemic period), increasing top management tournament incentives motivates the top executives work harder on inventory efficiency. In firms with larger bullwhip effect, top executive pay dispersion has a stronger positive impact on inventory management efficiency. However, when facing unprecedently uncertain situation such as COVID-19 pandemic, the incentive effect from tournament incentives is less pronounced. Practitioners can learn from these insights and induce more effort from top managers based on the findings of this study.

The rest of the paper proceeds as follows. Section 2 reviews relevant literature and develops hypotheses. Section 3 presents our sample, data, and the empirical models. We report data analyses and main results in Section 4. Section 5 concludes the paper with a discussion of theoretical and practical implications of the study.

2. Hypothesis Development

2.1 Top executive pay dispersion and inventory efficiency

Inventory management, an important business policy, is critical to improve firm operational and financial performance (Chuang et al, 2019; Barker et al, 2022). Prior studies have shown that reducing inventory holdings minimizes waste, thereby increasing efficiency (e.g.,

Callen et al., 2000; Womack et al., 2007). Along this line, Chen et al. (2005) show that manufacturing firms in the U.S. significantly reduced their inventories over the 1981-2000 period. They also find that a firm's inventory level is negatively associated with its long-term stock market performance.[1] In addition, Eroglu and Hofer (2011) document that the significance and shape of the relationship between inventory leanness and firm performance, as measured by return on sales and return on assets, vary across industries. Furthermore, they show that the effect of inventory leanness on firm performance is mostly positive and concave. Examining manufacturing firms' debt financing, Bendig et al. (2017) provide evidence that inventory leanness is positively associated with credit ratings in a concave relationship. Overall, these studies suggest that there is an optimal level of inventory. However, inventory efficiency decreases and the marginal effect of inventory leanness on financial performance becomes negative when inventory goes below the optimal level.[2] For instance, Feng et al. (2015) show that the lack of proper inventory management in inventory ordering, tracking, and evaluation has a direct impact on firm's operating performance. Therefore, both researchers and practitioners are interested in what incentives motivate top executives to achieve a higher level of inventory efficiency.

Tournament theory suggests that workers are provided with promotion-based incentives within organizations (Lazear and Rosen 1981; Green and Stokey 1983). In order to induce effort, firms tend to determine pay through the rank-order and increase the pay gap up the hierarchy. In a typical rank-order tournament system, only the best relative performer is promoted to the top level in the firm, i.e., the CEO position, and thus receives the highest compensation package.

[1] Specifically, they document that firms with abnormally high inventories have abnormally poor stock returns, while firms with slightly lower than average inventories perform best over time. In addition, firms with the lowest inventories have only ordinary returns.

[2] The optimal level of inventory depends on factors such as inventory carrying costs, shortage costs, and production technology.

Such a tournament provides top executives with incentives to work harder and adopt different strategies to increase their chances of promotion. Consistent with this, prior literature has documented a positive association between tournament incentives (proxied by the compensation dispersion among the CEO and vice presidents) and firm performance (measured by ROA and Tobin's Q as in Lee et al. 2008; Kale et al. 2009). Furthermore, tournament theory suggests that promotion-based incentives motivate top managers to engage in more risk-taking actions. Supporting this view, Kini and Williams (2012) find that greater pay dispersion between CEO and vice presidents is positively associated with riskier investment and financial policy choices (e.g., less industry diversity and higher leverage), resulting in greater firm riskiness (i.e., higher cash flow volatility and stock return volatility).

Managers have incentives to manage inventories efficiently because inventories represent not only an important asset but also a strategic resource. As discussed above, lower levels of inventory holdings indicate higher operational efficiency and is vital to supply chain management. Tournament literature suggests that greater pay dispersion among top executives has two effects on inventory efficiency. First, a tournament reward system motivates managers to exert more performance-increasing effort to improve their probability of promotion. As higher inventory efficiency improves operational as well as financial performance, top executives are likely to lower inventory holdings to the optimal level and meanwhile ensure the inventory level does not go below the optimal level, so that inventory leanness does not hurt firm performance. Second, as option-like tournament incentives motivate managers to take on greater risk (Kini and Williams 2012), they are willing to tolerate more operational risk. In the context of inventory management, it is likely that top executives, motivated by tournament incentives, are reluctant to increase inventory holdings, which represents a common way to

hedge operational risk. If managers choose to maintain inventories below the optimal level, the firm will not be able to increase flexibility and scalability to positively address changing market demand, neither can it utilize an inventory buffer to overcome supply chain disruptions. In other words, tournament incentives could discourage firms from carrying sufficient inventory and thus adversely affect their ability to cope with both supply and demand shocks. Hence, whereas managerial risk-taking behavior motivated by tournament incentives can improve a firm's inventory efficiency by boosting inventor leanness, it is likely to push it below the optimal level. This can expose firms to a significant amount of operational risk under extreme operating environments.

To summarize, tournament incentives of top executives motivate managers to engage in performance-improving and risk-inducing behavior, which leads to lower inventory levels. However, it is an empirical question whether tournament incentives can improve inventory efficiency and thus firm performance, given the positive and concave relationship between inventory leanness and firm performance. Accordingly, our first hypothesis is stated in null form.

H1*: Top executive pay dispersion is not associated with inventory efficiency.*

2.2 The impact of bullwhip effect

The bullwhip effect, widely found in business, exerts a significant impact on inventory efficiency. It describes a phenomenon in which the variability of order tends to be amplified as one moves from downstream to upstream in a supply chain (Lee et al. 1997). The bullwhip effect has dire consequences for each entity involved in supply chain including inefficient production and excessive inventory. Instead of examining this effect at the supply chain level, our paper investigates firm-level bullwhip effect to better understand the impact of top executive tournament incentives on inventory management practices. As the bullwhip effect may disrupt firms' daily

business transactions by enlarging the difference in variance between production and sales, it can lead to significant changes in operational performance within the firm.

Previous studies have investigated the relationship between the inventory level and the bullwhip effect. Theoretical papers (e.g., Tsay and Lovejoy 1999; Chen and Disney 2003) predict that when a firm faces demand with larger order variation, it will need to stock a greater level of inventory to maintain the same level of service, resulting in higher inventory cost. Based on a set of laboratory experiments, Croson et al. (2014) find that additional on-hand inventory is used as a managerial intervention to decrease the bullwhip effect by providing an inventory buffer. In addition, Shan et al (2014) empirically show that the bullwhip effect is positively associated with inventory days and inventory levels using Chinese data. Similarly, Mackelprang and Malhotra (2015) document that publicly traded U.S. firms facing an amplified bullwhip effect tend to possess higher level of inventory. In summary, prior literature suggests that the bullwhip effect prompts firms to hold increased levels of inventory as a precaution.

As higher inventory level resulting from a greater bullwhip effect harms inventory efficiency and operational performance, we predict that top executives motivated by stronger tournament incentives will respond more proactively in inventory management. This in turn mitigates the negative impact of the bullwhip effect. To achieve better firm performance, they are more likely to improve inventory efficiency when facing a greater bullwhip effect. Specifically, these managers can adopt strategies such as promoting coordination and information sharing between functions and implementing more accurate forecasting methods. Furthermore, as top executives are willing to hold lower inventory levels and bear more operational risk, they are less likely to sacrifice quick inventory turnover by maintaining safety stock to buffer against uncertainty in demand or deliveries. Put differently, we hypothesize that the impact of top

executive tournament incentives on inventory efficiency is more pronounced when there are more challenges in inventory management due to greater bullwhip effect. This leads to our second hypothesis:

H2*: The effect of top executive pay dispersion on inventory efficiency is stronger when firms are subject to a greater bullwhip effect.*

2.3 Impact of the COVID-19 pandemic

Ever since the World Health Organization (WHO) declared COVID-19 a pandemic in March 2020, it has had far-reaching consequences in the business world. Specifically, the pandemic has created significant uncertainty in all areas of business, resulting in unprecedented vulnerabilities in lead times and order quantities as well as severe demand fluctuations (Ivanov and Dolgui, 2021). This has unveiled an understudied area on managing business under extremal shocks of exogenous dynamics (Dolgui and Ivanov, 2021; Gupta, et al., 2022) including inventory management.

The pandemic is likely to prompt top managers to rethink inventory management and shift the focus from inventory efficiency to more critical issues such as dealing with shortages of inputs and disruptions of supply chains. For example, firms must decide if more inventories should be pre-positioned to overcome the material shortages during the pandemic. Consistent with this, an interview conducted by Rozhkov et al. (2022) during pandemic with top executives has revealed several critical inventory-related issues, which include suppliers being unable to fulfill demand, orders being cancelled, having difficulty to find new suppliers, and unprecedented longer lead time to fill the orders.

Besides an elevated level of uncertainty in supply, demand variability also brings more challenges to top management. For example, companies like Walmart have experienced panic

buying due to the COVID-19 pandemic for items ranging from hand sanitizer and toilet paper to hair products. Using a simulation-based analysis, Ivanov (2020) has confirmed the increased uncertainty level of supply and demand fluctuations during pandemic. To deal with the uncertainties coming from both supply and demand ends, some companies reported that they invested millions of dollars in inventory in anticipation of the havoc the pandemic was projected to create (Rozhkov et al., 2022).

In summary, the motivation to improve inventory efficiency coming from tournament incentives will be weakened when facing an unprecedent level of uncertainties and challenges. The leads to the third hypothesis:

H3*: During the COVID-19 pandemic, the effect of top executive pay dispersion on inventory efficiency is weakened regardless of an increased degree of the bullwhip effect.*

3. Methodology

3.1 Sample Selection

We test our hypotheses by conducting empirical analyses using data publicly available from SEC's EDGAR database. Specifically, our initial sample includes all U.S. manufacturing firms (with SIC codes between 2000 and 3999) available in both Compustat executive compensation and Compustat fundamentals annual databases from 1994 to 2021. To properly calculate top executive tournament incentives, we require that each firm-year observation in our sample has executive compensation data available for all top five paid executives and that the CEO has the highest amount of total compensation. We also require that data be available to calculate our measures of inventory efficiency, CEO equity compensation incentives, and important firm characteristics such as the volatility of sales revenue over a rolling window of 5 years for a given year. These procedures yield an unbalanced cross-sectional time-series panel with total of 12,669

firm-year observations for 1,153 different manufacturing firms. Panel A of Table 1 reports the number of observations per year from 1994 to 2021. The number of firms each year ranges from 301 (in 2021) to 548 (in 2008). Panel B of Table 1 presents the number of observations in each 2-digit SIC industry group. Untabulated results show that the industry composition of the sample is comparable to that of Compustat industrial annual files.

[Insert Table 1 about here]

3.2 Regression Models

We measure inventory efficiency in five different ways.[3] Our first measure is total inventory scaled by total assets (INVTO). The second measure is the difference between a firm's INVTO, and its 3-digit SIC industry mean value for a given year (ABNINVTO). We assume that ABNINVTO captures the abnormal portion of INVTO relative to the industry norm. The third measure, RESINVTO, is an estimate of the abnormal inventory level, calculated as the residual from the annual regressions of firms in the same 3-digit SIC industry group as in the following model:

$$\frac{INVT_{it}}{AT_{,t-1}} = \alpha_0 + \beta_1 \left(\frac{1}{AT_{it-1}}\right) + \beta_2 \left(\frac{REVT_{it}}{AT_{it-1}}\right) + \beta_3 \left(\frac{\Delta REVT_{it}}{AT_{it-1}}\right) + \varepsilon_{it-1},$$

where INVT is inventory, AT is total assets, and REVT is sales. Lower levels of INVTO, ABNINVTO, and RESINVTO indicate less stockpiling and more leanness, thus higher efficiency in inventory management. Our fourth measure is inventory turnover (INV_TURN), calculated as cost of goods sold over average inventory. Higher levels of inventory turnover are associated with stronger sales and more profitable operations. Our last measure captures inventory leanness

[3] We use the first three measures of inventory management efficiency following Hamm et al. (2022). Our fourth measure, inventory turnover, has been widely used by prior studies as a measure of inventory efficiency, such as Demeter and Matyusz (2011). Our fifth measure, inventory leanness, is measured following Bendig et al. (2017).

(INV_ELI), estimated as -1 times residuals from annual regressions of the natural logarithm of inventory on the natural logarithm of sales for each 4-digit SIC industry group. Greater values of INV_ELI indicate higher degrees of inventory leanness after controlling for industry and firm size.

We measure top executive tournament incentives empirically by top management pay dispersion, following Kale et al. (2009) and Kini and Williams (2012). Pay dispersion (PAY_DISP) is our independent variable of interest and is calculated as the natural log of the difference between the CEO's total compensation and the median total compensation of the next four highest paid executives. To examine the effect of top executive tournament incentives on firms' inventory management, we regress *current* inventory efficiency measures on *past* pay dispersion and controls including the bullwhip effect and other firm characteristics. Specifically, we perform OLS regressions with 2-digit SIC code industry and year effects, where standard errors are cluster-adjusted at the firm level.[4] Our regression model is as follows:

$$
\begin{aligned}
&INV\_EFFIC_t \\
&= \beta_0 + \beta_1 PAY\_DISP_{t-1} + \beta_2 CEO\_DELTA_{t-1} + \beta_3 CEO\_VEGA_{t-1} \\
&+ \beta_3 BULLWHIP_{t-1} + \beta_4 ROA_{t-1} + \beta_5 LEVER_{t-1} + \beta_6 Q_{t-1} + \beta_7 PPE_{t-1} \\
&+ \beta_8 SALES\_GROW_{t-1} + \beta_9 SALES\_VOL_{t-1} + \beta_{10} LNMV_{t-1} + Other\ Controls \\
&+ Industry\ Indicators + Year\ Indicators + \varepsilon_t
\end{aligned}
$$

where INV_EFFIC is a firm's inventory efficiency measured by INVTO, ABNINVTO, RESINVTO, INV_TURN, and INV_ELI. For measures INVTO, ABNINVTO, RESINVTO, a negative coefficient on PAY_DISP would be consistent with the notion that greater top executive tournament incentives lead to a lower inventory level and thus more efficient inventory management. For measures INV_TURN and INV_ELI, a positive coefficient on PAY_DISP

[4] Following Petersen (2009), our statistics are based on clustered standard errors by firm, which account for the residual dependence created by the firm effect and thus are unbiased.

would indicate that greater top executive tournament incentives lead to higher inventory turnover and leaner inventory.

We also control for the effect of executive equity compensation incentives on inventory management, following Kale et al. (2009) and Kini and Williams (2012). As in Core and Guay (2002), we calculate CEO pay-performance sensitivity as the change in the dollar value of CEO wealth for a 1% change in stock price (CEO_DELTA), and CEO pay-risk sensitivity as the change in the dollar value of CEO wealth for a 0.01 change in stock return volatility (CEO_VEGA).[5] We predict that both CEO_DELTA and CEO_VEGA are positively related to inventory efficiency measures.

A particularly important control variable in our regression is BULLWHIP, which measures the extent to which a firm is subject to the bullwhip effect. Following Shan et al. (2014), we estimate BULLWHIP by the ratio of volatility of production over the volatility of demand. We first calculate production as a firm's cost of goods sold for a given year plus its inventory level at the end of the year minus its inventory level at the beginning of the year. Next we calculate the volatility of production as the standard deviation of a firm's production over the 5-year period prior to the current year. Finally. we use a firm's cost of goods sold for a given year to proxy for its product demand and calculate the volatility of demand as the standard deviation of a firm's cost of goods sold over the 5-year period prior to the current year. A higher level of BULLWHIP indicates that there is a greater degree in amplifications of demand volatility along the supply chain, leading to less efficient inventory management. Therefore, we predict that the coefficient on BULLWHIP is positive when our dependent variables are INVTO, ABNINVTO, and

[5] Untabulated results show that controlling for delta and vega of the top 5 executive officers does not change our inferences.

RESINVTO, and the coefficient on BULLWHIP is negative when our dependent variables are INV_TURN and INV_ELI.

Following Feng et al. (2015) and Hamm et al. (2022), we also control for classic firm performance and risk measures, i.e., the return-on-asset ratio (ROA), Tobin's Q (Q), annual sales growth (SALES_GROW), and annual sales volatility (SALES_VOL). Prior studies have documented that the above four firm characteristics are related to product demand and thus inventory levels. Furthermore, we control for firm investing and financing characteristics such as leverage (LEVER), asset tangibility (PPE), and market capitalization (LNMV). Prior literature has shown that they affect the stability of inventory levels. In addition, we include capital expenditure (CAPEX) in regressions of INVTO, ABNINVTO, and RESINVTO, and we include the gross margin ratio (GM_RATIO) and an indicator for foreign sales (FSALES) in regressions of INV_TURN and INV_ELI. The variable definitions are summarized in the Appendix. Each variable is winsorized to the top and bottom 1% of its distribution to reduce the effects of extreme observations.

3.3 Descriptive Statistics

The descriptive statistics of the variables used in our tests appear in Panel A of Table 2. The median of INVTO is 0.121, which is lower than the median of 0.17 reported in Hamm et al. (2020). This is due to the fact that our sample firms are much larger than theirs, and firm size is documented to be negatively related to the ratio of inventory to total assets. The median of INV_TRUN is 4.308 and is comparable with the median reported in Feng et al. (2015). The median of PAY_DISP is 7.458 and is in line with the mean of 7.257 reported in Kini and Williams (2012). Finally, the estimated delta and vega are on average $628 and $132 before the log transformation, which are comparable to those reported in Kini and Williams (2012). Our sample median of

BULLWHIP is 1.150 and the 25th percentile is 1.001, indicating that we have 75% of firm-years with bullwhip ratios greater than 1 and thus suffering from the amplification of demand variability. Furthermore, the sample means of our other control variables are comparable to those in the prior literature.

[Insert Table 2 about here]

Panel B of Table 2 reports the Pearson correlations between inventory efficiency measures, pay dispersion, and control variables. It shows that INVTO, ABNINVTO, and RESINVTO are negatively related to PAY_DISP. Although the simple linear correlation between INV_TRUN and PAY_DISP is not significantly different from zero, the correlation becomes significantly positive (coefficient = 0.063) after controlling for GM_RATIO.[6] Similarly, the simple linear correlation between INV_ELI and PAY_DISP is significantly negative, but the correlation becomes significantly positive (coefficient = 0.026) after controlling for ROA and LNMV. These correlations suggest that there is a positive association between a firm's inventory management efficiency and top management pay dispersion. However, given the significant correlations between the control variables and pay dispersion, a multivariate analysis is needed to confirm these preliminary univariate findings. As expected, CEO_DELTA and CEO_VEGA are in general positively related to inventory efficiency measures. Furthermore, there are high pairwise correlations between the five measures of inventory management measures with the absolute values of coefficients ranging from 0.232 to 0.818. Consistent with Kale et al. (2009) and Kini and Williams (2012), the pay dispersion variable is positively related to ROA, Tobin's Q, leverage, firm market value, gross margin ratio, and the occurrence of foreign sales, and it is negatively

[6] It is likely that the statistically insignificant simple correlation is driven by the fact that the highly negative correlation between INV_TURN and GM_RATIO (coefficient = 0.359) obscures the positive correlation between INV_TURN and PAY_DISP.

related to capital intensity, capital expenditures, and sales volatility. As in Feng et al. (2015) and Hamm et al. (2022), inventory efficiency is positively related to leverage, Tobin's Q, sales growth, firm market value, gross margin ratio, and the presence of foreign sales, and it is negatively related to the bullwhip effect, capital intensity, sales volatility, and capital expenditures. Finally, the correlations among other controls are also in line with the prior literature.

4. Results

4.1 Inventory Management Efficiency and Top Executive Tournament Incentives

We hypothesize that top executives in manufacturing firms with greater tournament incentives, as measured by top executive pay dispersion, are more likely to decrease inventory levels, increase inventory turnover, and inventory leanness due to the performance-increasing effect as well as the risk-increasing effect. These actions promote inventory management efficiency as long as the inventory holding does not go below the optimal level. To test this hypothesis, we run OLS regressions of various measures of inventory management on top executive pay dispersion and control variables. The regression results of Equation (1) are reported in Table 3. We find that the coefficients on PAY_DISP in the first three columns are statistically significant and negative at the .01 level. For example, the coefficient on PAY_DISP is -0.004 in column (2), which suggests that firms with larger pay dispersion among top executives are likely to maintain lower inventory levels compared to the industry norm. The magnitude of the coefficient shows that the effect of pay dispersion on firm inventory management is economically significant. Specifically, a one standard deviation increase (1.213) in a firm's PAY_DISP is associated with a decrease in the firm's abnormal inventory level by 15.6%, holding other variables in the model constant. Furthermore, the coefficients on PAY_DISP in regressions of INV_TURN and INV_ELI are significantly positive at the .01 level and at the .05 level, respectively, in columns

(4) and (5). This is consistent with the notion that firms with higher pay dispersion among top executives are more likely to have higher inventory turnover and greater inventory leanness. Taken together, the results in Table 3 support H1, and indicate that larger pay dispersion among top executives motivate them to improve firm performance and tolerate more operational risk in their inventory-related policy decisions. As a result, firms that provide their top executive with more tournament incentives benefit from this compensation practice by having more efficient inventory management.

[Insert Table 3 about here]

Interestingly, whereas the coefficients on CEO_DELTA delta are not significantly different from zero, those on CEO_VEGA are significantly negative in columns (1)-(3), showing that a higher pay-risk sensitivity is associated with a lower level of inventories. This indicates that top executives with more risk-taking incentives are less likely to maintain safety stock to buffer against uncertainty. We also find that the coefficients on BULLWIP are significantly positive in columns (1)-(3) and significantly negative in Columns (4) and (5). This is consistent with prior evidence that firms suffering from the bullwhip effect are less likely to have efficient inventory management. In general, the estimated coefficients on other control variables are consistent with prior studies. In columns (1)-(3), the coefficients on ROA, PPE, and SALES_VOL are significantly positive, and the coefficients on LEVER, Q, LNMV, and CAPEX are significantly negative. These results show that firms with worse accounting performance, less tangible assets, and less sales volatility and firms with higher leverage, better stock performance, larger, and more capital expenditures are more likely to have lower levels of inventory. In columns (4) and (5), the coefficients on Q, PPE, and SALES_VOL are significantly positive, and the coefficients on

GM_RATIO and FSALES are significantly negative. These results again are in line with those documented in the prior literature,

4.2 The Interaction between Top Executive Tournament Incentives and the Bullwhip Effect

Our second hypothesis (H2) predicts the effect of top executive tournament incentives on inventory management efficiency is stronger when a firm suffers more from the bullwhip effect. To test H2, we define H_BULLWHIP as 1 if a firm-year observation has a BULLWHIP value that is 1.21 or higher, 0 otherwise. We use 1.21 as the cutoff to indicate observations suffering a greater bullwhip effect because it is the 75th percentile of the variable BULLWHIP in our full sample. We then interact H_BULLWHIP with PAY_DISP, and we expect that a negative coefficient on the interaction term would indicate that the effect of tournament incentives is stronger when firms are subject to a higher degree of the bullwhip effect. The regression results are reported in Table 4. As expected, the coefficients on PAY_DISP* H_BULLWHIP are significantly negative in columns (1)-(3) and are significantly positive in columns (4)-(5). For instance, the coefficient on PAY_DISP* H_BULLWHIP is -0.001 (significant at the 1% level) and the coefficient on PAY_DISP is -0.004 (significant at the 5% level) in column (2). This means that, if there is a one standard deviation increase (1.213) in the PAY_DISP of a firm having a larger bullwhip effect, there is a decrease in the firm's abnormal inventory level by 19.6%, holding other variables in the model constant. To conclude, our evidence in Table 4 lends support to H2, and shows that the effect of top executive tournament incentives on inventory management efficiency is more pronounced for firms subject to stronger bullwhip effect.

[Insert Table 4 about here]

4.3 The Impact of the COVID-19 pandemic

Our third hypothesis (H3) states that the effect of top executive tournament incentives on inventory efficiency is weaker during the COVID-19 pandemic even if a firm is subject to a greater bullwhip effect. To test H3, we first identify 12,005 observations in the pre-pandemic period (firm-years with fiscal year ending before or in February 2020) and 664 observations in the post pandemic period (firm-years with fiscal year ending in or after March 2020). In Panel A of Table 5, we compare the means of our variables in the pre-pandemic versus post-pandemic periods. Consistent with the prior literature, although total inventory levels (INVTO) decreased, abnormal inventory levels (ABNINVTO) increased and inventory turnover decreased, and more importantly, the bullwhip effect worsened in the post-pandemic period. As for compensation contracting, tournament incentives (PAY_DISP) and CEO pay-performance sensitivity (CEO_DELTA) are stronger, whereas CEO pay-risk sensitivity (CEO_VEGA) are weaker in the post-pandemic period.

[Insert Table 5 about here]

We then run regressions separately on two subsamples after we split our sample period into pre-pandemic and post-pandemic. The regression results for the subsample in the pre-pandemic period are reported in Panels B and C, and the regression results for the subsample in the post-pandemic period are reported in Panels D and E.[7] The results in Panels B and C of Table 5 are qualitatively similar to those reported in Tables 3 and 4. Specifically, the coefficients on PAY_DISP are significantly negative in columns (1)-(3) and significantly positive in columns (4) and (5) in Panel B. Furthermore, the coefficients on PAY_DISP*H_BULLWHIP are significantly negative in columns (1)-(3) and significantly positive in columns (4) and (5) in Panel C. Also, consistent with our previous results, the coefficients on BULLWHIP are significantly positive in

[7] Alternatively, we interact an indicator variable PANDEMIC with PAY_DISP and PAY_DISP*H_BULLWHIP and run the regressions for the full sample. Our inferences from the unreported results remain unchanged.

columns (1)-(3) and significantly negative in columns (4) and (5) in Panels B and C. As predicted, the pre-pandemic results indicate that tournament incentives among top executives motivate them to promote inventory efficiency, especially when firms have a bigger bullwhip problem. In contrast, the confounding effect of externalities unique to the pandemic has completely changed the results in Panels D and E, when we examine the subsample in the post-pandemic period. The bullwhip effect still leads to higher inventory levels and lower inventory turnover and leanness. However, pay dispersion among top executives does not show any effect on inventory management anymore. In addition, coefficients on PAY_DISP*H_BULLWHIP are not significant in the three inventory level regressions. The coefficient on PAY_DISP*H_BULLWHIP is significantly positive in the inventory turnover regression, and it is partially significant in the inventory leanness regression. This shows that tournament incentives only have limited effect on inventory management in the post-pandemic period, as top executives are coping with the increased demand variability and supply uncertainty and supports H3.

5. Conclusion

This paper investigates the impact of top executive tournament incentives on inventory efficiency. Inventory efficiency, as an important indicator of operational performance, has been pursued by all types of firms and their top executives. Therefore, factors determining inventory efficiency are of wide interest, and warrant thorough investigation. We provide evidence supporting a positive relationship between tournament incentives (proxied by pay dispersion among top executives) and inventory efficiency. Our findings also reveal that the effect of top executive tournament incentives on inventory management efficiency is stronger for firms subject to stronger bullwhip effect. Finally, when separating the sample into pre-pandemic and post-

pandemic, we find that pay dispersion among top executives has no effect on inventory efficiency in the post-pandemic era.

To our best knowledge, this is the first empirical study to investigate the impact on inventory efficiency from the perspective of top management tournament incentives. Our findings contribute to the research on inventory management by showing tournament incentive among top executives is an important determinant of inventory efficiency. In addition, taking into account the roles played by bullwhip effect and COVID-19 pandemic, this study enriches the research on consequences of top management tournament incentives.

Our results also provide practical implications and guidance for practitioners. In business environment with fewer uncertainties (i.e., prior to pandemic), increasing top management tournament incentives motivates top executives to work harder and improve inventory efficiency. Furthermore, with respect to firms with greater bullwhip effect, top executive pay dispersion has a stronger positive impact on inventory management efficiency. However, when facing unprecedently uncertainties such as COVID-19 pandemic, tournament incentives will have a lessened effect. Practitioners might learn from these insights and induce more effort from top managers accordingly.

Appendix: Variable Definition

Variable	Definitions
INVTO	Total inventory scaled by total assets.
ABNINVTO	Abnormal INVTO relative to its 3-digit SIC industry-year average.
RESINVTO	Residual from 3-digit SIC industry-year regressions of $\frac{INVT_{it}}{AT_{,t-1}} = \alpha_0 + \beta_1\left(\frac{1}{AT_{it-1}}\right) + \beta_2\left(\frac{REVT_{it}}{AT_{it-1}}\right) + \beta_3\left(\frac{\Delta REVT_{it}}{AT_{it-1}}\right) + \varepsilon_{it-1}$.
INV_TURN	Cost of goods sold over average inventory
	-1 times Residual from 4-digit SIC industry-year regressions of $Ln(INVT_{it}) = \alpha_0 + \beta_1 Ln(REVT_{it}) + \varepsilon_{it}$.
PAY_DISP	Natural log of difference between CEO total compensation and the median of total compensation paid to the next four highest paid top executives at the end of the current fiscal year
CEO_DELTA	Natural log of pay-performance sensitivity of the firm's CEO as in Core and Guay (2002)
CEO_VEGA	Natural log of pay-risk sensitivity of the firm's CEO as in Core and Guay (2002)
BULLWHIP	The ratio of volatility of production over the volatility of demand. Production is calculated as a firm's cost of goods sold for a given year plus its inventory level at the end of the year minus its inventory level at the end of the year. Demand is proxied by cost of goods sold.
ROA	Net income before extraordinary items divided by total assets.
LEVER	Book value of debt scaled by total assets.
Q	Tobin's Q, measured as the ratio of the market value of total assets to the book value of total assets.
PPE	Gross PPE scaled by total assets.
SALES_GROW	Growth rate in total sales revenue.
SALES_VOL	Standard deviation of sales scaled by total assets over the past five-year period.
LNMV	Natural log of the market value of equity.
CAPEX	Capital expenditure scaled by total assets.
GM_RATIO	The gross profit ratio measured as sales minus cost of sales divided by sales.
FSALES	An indicator variable equal to 1 if the firm reports foreign sales, and 0 otherwise.

Table 1 Distribution of Sample Observations

Panel A: Annual Distribution

Year	# of Observations
1994	361
1995	438
1996	471
1997	475
1998	467
1999	476
2000	481
2001	470
2002	453
2003	476
2004	499
2005	489
2006	489
2007	466
2008	548
2009	538
2010	523
2011	498
2012	476
2013	459
2014	456
2015	443
2016	419
2017	414
2018	388
2019	332
2020	363
2021	301
Total	12,669

Panel B: Industrial Distribution

2-digit SIC	Description	# of Observations
20	Food And Kindred Products	722
22	Textile Mill Products	59
23	Apparel And Other Finished Products Made from Fabrics and Similar Materials	255
24	Lumber and Wood Products, Except Furniture	126
25	Furniture and Fixtures	123
26	Paper and Allied Products	351
27	Printing, Publishing, and Allied Industries	205
28	Chemicals and Allied Products	2,389
29	Petroleum Refining and Related Industries	134
30	Rubber and Miscellaneous Plastics Products	164
31	Leather and Leather Products	77
32	Stone, Clay, Glass, and Concrete Products	79
33	Primary Metal Industries	489
34	Fabricated Metal Products, Except Machinery and Transportation Equipment	308
35	Industrial and Commercial Machinery and Computer Equipment	1,962
36	Electronic and Other Electrical Equipment and Components, Except Computer Equipment	2,396
37	Transportation Equipment	724
38	Measuring, Analyzing, and Controlling Instruments; Photographic, Medical and Optical Goods; Watches and Clocks	1,882
39	Miscellaneous Manufacturing Industries	224
Total		12,669

Table 2 Summary Statistics and Correlations

Panel A: Summary Statistics (N=12,669)

Variable	Mean	Standard Deviation	Lower Quartile	Median	Upper Quartile
INVTO	0.134	0.079	0.076	0.121	0.175
ABNINVTO	-0.031	0.070	-0.080	-0.040	0.006
RESINVTO	-0.022	0.064	-0.062	-0.027	0.011
INV_TURN	5.504	4.798	2.899	4.308	6.237
INV_ELI	-0.005	0.453	-0.301	-0.015	0.258
PAY_DISP	7.458	1.213	6.684	7.565	8.347
CEO_DELTA	5.270	1.457	4.340	5.252	6.220
CEO_VEGA	3.653	1.798	2.643	3.875	4.912
BULLWHIP	1.150	0.287	1.001	1.082	1.216
ROA	0.053	0.106	0.017	0.060	0.103
LEVER	0.210	0.167	0.054	0.204	0.320
Q	2.150	1.408	1.294	1.716	2.443
PPE	0.481	0.309	0.249	0.406	0.645
SALES_GROW	0.097	0.229	-0.014	0.070	0.170
SALES_VOL	0.209	0.196	0.084	0.146	0.260
LNMV	7.447	1.644	6.306	7.296	8.437
CAPEX	0.044	0.034	0.021	0.035	0.057
GM_RATIO	0.421	0.195	0.276	0.400	0.556
FSALES	0.424	0.494	0	0	1

Variable definitions are available in the Appendix.

Penal B: Pearson Correlations (N=14,095)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	**INVTO**	1.000																	
2	**ABNINVTO**	**0.816**	1.000																
3	**RESINVTO**	**0.605**	**0.741**	1.000															
4	**INV_TURN**	**-0.303**	**-0.222**	**-0.355**	1.000														
5	**INV_ELI**	**-0.427**	**-0.493**	**-0.644**	**0.494**	1.000													
6	**PAY_DISP**	**-0.245**	**-0.188**	**-0.125**	0.009	**-0.022**	1.000												
7	**CEO_DELTA**	**-0.204**	**-0.186**	**-0.132**	**-0.010**	-0.001	**0.418**	1.000											
8	**CEO_VEGA**	**-0.245**	**-0.207**	**-0.130**	**-0.032**	**-0.026**	**0.459**	**0.528**	1.000										
9	**BULLWHIP**	**0.069**	**0.083**	**0.123**	**-0.187**	**-0.139**	-0.002	-0.015	**0.022**	1.000									
10	**ROA**	**-0.020**	**-0.057**	**-0.118**	**-0.024**	**0.148**	**0.127**	**0.301**	**0.100**	**-0.028**	1.000								
11	**LEVER**	**-0.036**	**-0.029**	**-0.030**	**0.064**	**-0.037**	**0.188**	**0.024**	**0.081**	**-0.028**	**-0.127**	1.000							
12	**Q**	**-0.176**	**-0.120**	**-0.145**	**-0.065**	**0.147**	**0.131**	**0.408**	**0.133**	**0.028**	**0.358**	**-0.187**	1.000						
13	**PPE**	**0.085**	**0.115**	0.008	**0.231**	**0.070**	**-0.100**	**-0.154**	**-0.099**	**-0.051**	-0.013	**0.145**	**-0.203**	1.000					
14	**SALES_GROW**	**-0.036**	**-0.051**	**-0.055**	**0.013**	**0.058**	0.008	**0.164**	**0.020**	0.011	**0.152**	**-0.064**	**0.285**	**-0.173**	1.000				
15	**SALES_VOL**	**0.090**	**0.070**	**-0.023**	**0.104**	**0.086**	**-0.175**	**-0.083**	**-0.122**	**-0.030**	**-0.085**	**-0.096**	**0.039**	**-0.124**	**0.150**	1.000			
16	**LNMV**	**-0.315**	**-0.239**	**-0.166**	**-0.005**	**-0.039**	**0.662**	**0.622**	**0.492**	0.002	**0.307**	**0.131**	**0.347**	**-0.080**	**0.052**	**-0.246**	1.000		
17	**CAPEX**	**0.024**	**-0.021**	**-0.064**	**0.131**	**0.060**	**-0.112**	**0.042**	**-0.058**	**-0.089**	**0.072**	**-0.053**	**0.085**	**0.488**	**0.079**	**0.054**	**-0.022**	1.000	
18	**GM_RATIO**	**-0.373**	**-0.263**	**-0.133**	**-0.359**	**0.057**	**0.128**	**0.267**	**0.223**	**0.160**	**0.217**	**-0.185**	**0.441**	**-0.387**	**0.116**	**-0.113**	**0.254**	**-0.087**	1.000
19	**FSALES**	**-0.023**	**0.054**	**0.050**	**-0.110**	**-0.085**	**0.119**	-0.003	**0.037**	**0.020**	**-0.020**	-0.008	**-0.026**	0.013	**-0.060**	**-0.116**	**0.104**	**-0.069**	**0.041**

Variable definitions are available in the Appendix. Bolded coefficients are significant at the 5% level.

Table 3

VARIABLES	(1) INVTO	(2) ABNINVTO	(3) RESINVTO	(4) INV_TURN	(5) INV_ELI
PAY_DISP	-0.003*	-0.004***	-0.003***	0.221***	0.017**
	(-1.88)	(-2.97)	(-2.63)	(2.81)	(2.01)
CEO_DELTA	0.002	0.002	0.002	0.014	-0.012
	(1.37)	(1.12)	(1.53)	(0.16)	(-1.32)
CEO_VEGA	-0.004***	-0.003***	-0.003***	-0.009	0.008
	(-3.62)	(-3.20)	(-3.26)	(-0.14)	(1.21)
BULLWHIP	0.021***	0.020***	0.025***	-1.744***	-0.224***
	(5.38)	(5.50)	(7.61)	(-9.49)	(-8.75)
ROA	0.054***	0.018	-0.032***	-0.695	0.535***
	(3.70)	(1.41)	(-3.03)	(-0.83)	(6.06)
LEVER	-0.006	-0.020*	-0.019**	-0.795	0.061
	(-0.56)	(-1.93)	(-2.24)	(-1.21)	(0.93)
Q	-0.005***	-0.003**	-0.005***	0.298***	0.051***
	(-2.90)	(-2.03)	(-4.41)	(3.18)	(5.51)
PPE	0.024***	0.021***	-0.006	1.496***	0.183***
	(3.50)	(3.19)	(-1.11)	(2.71)	(3.42)
SALES_GROW	-0.005	-0.005	-0.002	0.724***	0.025
	(-1.45)	(-1.48)	(-0.50)	(2.86)	(1.01)
SALES_VOL	0.015**	0.018**	-0.017***	1.082**	0.162***
	(1.97)	(2.41)	(-2.72)	(2.36)	(3.61)
LNMV	-0.009***	-0.008***	-0.004***	0.076	-0.047***
	(-6.49)	(-6.27)	(-3.12)	(0.82)	(-4.66)
GM_RATIO				-7.672***	0.387***
				(-8.24)	(4.39)
FSALES				-0.622***	-0.062***
				(-3.42)	(-2.90)
CAPEX	-0.141***	-0.084**	-0.017		
	(-3.17)	(-2.04)	(-0.48)		
Constant	0.224***	0.031*	0.019	8.761***	0.088
	(12.99)	(1.95)	(1.35)	(7.41)	(0.73)
Observations	12,669	12,669	12,669	12,669	12,669
Adjusted R-squared	0.257	0.142	0.089	0.343	0.117
Year & Industry FE	YES	YES	YES	YES	YES

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4

VARIABLES	(1) INVTO	(2) ABNINVTO	(3) RESINVTO	(4) INV_TURN	(5) INV_ELI
PAY_DISP	-0.002	-0.004**	-0.002*	0.112	0.005
	(-1.41)	(-2.54)	(-1.75)	(1.45)	(0.60)
CEO_DELTA	0.002	0.002	0.002	0.008	-0.013
	(1.40)	(1.15)	(1.58)	(0.09)	(-1.41)
CEO_VEGA	-0.004***	-0.003***	-0.003***	-0.010	0.008
	(-3.63)	(-3.21)	(-3.28)	(-0.15)	(1.21)
BULLWHIP	0.013***	0.013***	0.013***	-0.578***	-0.094***
	(3.05)	(3.24)	(3.59)	(-3.30)	(-3.22)
PAY_DISP*H_BULLWHIP	-0.001***	-0.001***	-0.002***	0.147***	0.016***
	(-3.29)	(-2.97)	(-5.56)	(9.69)	(7.60)
ROA	0.054***	0.019	-0.031***	-0.805	0.523***
	(3.76)	(1.45)	(-2.95)	(-0.96)	(5.96)
LEVER	-0.006	-0.019*	-0.019**	-0.824	0.058
	(-0.53)	(-1.90)	(-2.19)	(-1.26)	(0.89)
Q	-0.005***	-0.003**	-0.005***	0.296***	0.051***
	(-2.93)	(-2.05)	(-4.48)	(3.18)	(5.53)
PPE	0.024***	0.021***	-0.007	1.518***	0.186***
	(3.50)	(3.19)	(-1.16)	(2.77)	(3.50)
SALES_GROW	-0.005	-0.005	-0.002	0.734***	0.026
	(-1.47)	(-1.50)	(-0.53)	(2.92)	(1.06)
SALES_VOL	0.016**	0.018**	-0.016***	1.037**	0.156***
	(2.02)	(2.46)	(-2.64)	(2.27)	(3.52)
LNMV	-0.009***	-0.008***	-0.004***	0.076	-0.047***
	(-6.51)	(-6.29)	(-3.15)	(0.83)	(-4.69)
GM_RATIO				-7.516***	0.404***
				(-8.13)	(4.62)
FSALES				-0.609***	-0.060***
				(-3.37)	(-2.85)
CAPEX	-0.139***	-0.082**	-0.014		
	(-3.13)	(-2.00)	(-0.39)		
Constant	0.233***	0.039**	0.033**	7.376***	-0.067
	(13.36)	(2.40)	(2.27)	(6.15)	(-0.55)
Observations	12,669	12,669	12,669	12,669	12,669
Adjusted R-squared	0.258	0.143	0.092	0.347	0.124
Year & Industry FE	YES	YES	YES	YES	YES

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5

Panel A Univariate analysis

Variables	Pre-Pandemic	Post-Pandemic	Difference	p-value
INVTO	0.135	0.116	0.019	***
ABNINVTO	-0.031	-0.019	-0.012	***
RESINVTO	-0.021	-0.023	0.001	
INV_TURN	5.563	4.435	1.127	***
INV_ELI	-0.004	-0.014	0.010	
PAY_DISP	7.420	8.152	-0.732	***
CEO_DELTA	5.259	5.492	-0.233	***
CEO_VEGA	3.670	3.265	0.405	***
BULLWHIP	1.148	1.182	-0.034	***
ROA	0.053	0.060	-0.007	*
LEVER	0.206	0.275	-0.068	***
Q	2.127	2.540	-0.413	***
PPE	0.484	0.430	0.054	***
SALES_GROW	0.101	0.018	0.083	***
SALES_VOL	0.213	0.146	0.067	***
LNMV	7.399	8.331	-0.932	***
CAPEX	0.045	0.030	0.015	***
GM_RATIO	0.419	0.463	-0.044	***
FSALES	0.417	0.545	-0.128	***

*** p<0.01, ** p<0.05, * p<0.1

Panel B Pre-Pandemic

VARIABLES	(1) INVTO	(2) ABNINVTO	(3) RESINVTO	(4) INV_TURN	(5) INV_ELI
PAY_DISP	-0.003*	-0.004***	-0.003**	0.220***	0.018**
	(-1.82)	(-2.97)	(-2.53)	(2.75)	(2.10)
CEO_DELTA	0.002	0.002	0.002	0.008	-0.014
	(1.44)	(1.21)	(1.48)	(0.09)	(-1.48)
CEO_VEGA	-0.005***	-0.004***	-0.003***	0.012	0.009
	(-3.86)	(-3.33)	(-3.42)	(0.18)	(1.35)
BULLWHIP	0.022***	0.020***	0.025***	-1.769***	-0.219***
	(5.22)	(5.30)	(7.12)	(-9.25)	(-8.37)
ROA	0.050***	0.016	-0.031***	-0.713	0.536***
	(3.39)	(1.18)	(-2.87)	(-0.82)	(5.97)
LEVER	-0.006	-0.019*	-0.019**	-0.849	0.062
	(-0.51)	(-1.85)	(-2.14)	(-1.22)	(0.92)
Q	-0.005***	-0.003**	-0.005***	0.325***	0.053***
	(-3.18)	(-2.35)	(-4.40)	(3.27)	(5.50)
PPE	0.023***	0.021***	-0.006	1.501***	0.180***
	(3.35)	(3.15)	(-1.12)	(2.64)	(3.35)
SALES_GROW	-0.005	-0.005	-0.001	0.730***	0.024
	(-1.40)	(-1.44)	(-0.38)	(2.71)	(0.96)
SALES_VOL	0.015**	0.017**	-0.017***	1.048**	0.157***
	(1.96)	(2.35)	(-2.68)	(2.24)	(3.49)
LNMV	-0.009***	-0.008***	-0.004***	0.059	-0.048***
	(-6.19)	(-6.01)	(-3.04)	(0.65)	(-4.76)
GM_RATIO				-7.866***	0.378***
				(-8.38)	(4.29)
FSALES				-0.597***	-0.057***
				(-3.28)	(-2.67)
CAPEX	-0.135***	-0.080*	-0.017		
	(-2.99)	(-1.92)	(-0.48)		
Constant	0.222***	0.031*	0.020	8.966***	0.091
	(12.83)	(1.96)	(1.40)	(7.39)	(0.76)
Observations	12,005	12,005	12,005	12,005	12,005
Adjusted R-squared	0.259	0.141	0.089	0.347	0.116
Year & Industry FE	YES	YES	YES	YES	YES

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Panel C Pre-Pandemic Interaction

VARIABLES	(1) INVTO	(2) ABNINVTO	(3) RESINVTO	(4) INV_TURN	(5) INV_ELI
PAY_DISP	-0.002	-0.004**	-0.002	0.109	0.006
	(-1.35)	(-2.51)	(-1.60)	(1.38)	(0.67)
CEO_DELTA	0.002	0.002	0.002	0.000	-0.014
	(1.48)	(1.24)	(1.55)	(0.00)	(-1.59)
CEO_VEGA	-0.005***	-0.004***	-0.003***	0.010	0.009
	(-3.87)	(-3.33)	(-3.44)	(0.16)	(1.34)
BULLWHIP	0.014***	0.013***	0.012***	-0.574***	-0.085***
	(2.96)	(3.05)	(3.15)	(-3.11)	(-2.84)
PAY_DISP*H_BULLWHIP	-0.001***	-0.001***	-0.002***	0.150***	0.017***
	(-3.11)	(-2.92)	(-5.57)	(9.52)	(7.43)
ROA	0.051***	0.016	-0.030***	-0.817	0.524***
	(3.44)	(1.22)	(-2.80)	(-0.94)	(5.88)
LEVER	-0.005	-0.019*	-0.018**	-0.883	0.058
	(-0.47)	(-1.82)	(-2.08)	(-1.28)	(0.87)
Q	-0.005***	-0.003**	-0.006***	0.323***	0.053***
	(-3.20)	(-2.38)	(-4.46)	(3.27)	(5.51)
PPE	0.023***	0.020***	-0.007	1.522***	0.182***
	(3.34)	(3.15)	(-1.15)	(2.69)	(3.42)
SALES_GROW	-0.005	-0.005	-0.001	0.742***	0.026
	(-1.43)	(-1.46)	(-0.42)	(2.78)	(1.02)
SALES_VOL	0.016**	0.018**	-0.016***	1.005**	0.152***
	(2.01)	(2.40)	(-2.60)	(2.16)	(3.40)
LNMV	-0.009***	-0.008***	-0.004***	0.060	-0.048***
	(-6.21)	(-6.03)	(-3.07)	(0.67)	(-4.78)
GM_RATIO				-7.694***	0.397***
				(-8.24)	(4.53)
FSALES				-0.583***	-0.055***
				(-3.23)	(-2.61)
CAPEX	-0.133***	-0.078*	-0.014		
	(-2.96)	(-1.89)	(-0.40)		
Constant	0.232***	0.040**	0.034**	7.553***	-0.067
	(13.16)	(2.42)	(2.35)	(6.16)	(-0.54)
Observations	12,005	12,005	12,005	12,005	12,005
Adjusted R-squared	0.260	0.142	0.092	0.352	0.124
Year & Industry FE	YES	YES	YES	YES	YES

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Panel D Post-Pandemic

VARIABLES	(1) INVTO	(2) ABNINVTO	(3) RESINVTO	(4) INV_TURN	(5) INV_ELI
PAY_DISP	-0.004	-0.003	-0.003	0.018	-0.010
	(-0.92)	(-0.69)	(-0.86)	(0.11)	(-0.37)
CEO_DELTA	0.001	0.000	0.003	0.082	0.014
	(0.33)	(0.12)	(0.92)	(0.44)	(0.53)
CEO_VEGA	-0.001	-0.001	-0.002	-0.194	-0.001
	(-0.50)	(-0.61)	(-0.98)	(-1.30)	(-0.08)
BULLWHIP	0.017**	0.017**	0.028***	-1.642***	-0.312***
	(1.97)	(1.99)	(3.81)	(-4.24)	(-4.42)
ROA	0.127***	0.104***	-0.050	-0.256	0.592**
	(3.51)	(2.98)	(-1.52)	(-0.17)	(2.20)
LEVER	-0.011	-0.026	-0.030	-0.044	0.071
	(-0.53)	(-1.22)	(-1.60)	(-0.04)	(0.42)
Q	0.001	0.002	-0.003*	0.050	0.031*
	(0.32)	(0.72)	(-1.65)	(0.44)	(1.75)
PPE	0.048***	0.031*	-0.005	1.142	0.282**
	(2.65)	(1.80)	(-0.36)	(1.63)	(2.17)
SALES_GROW	0.009	0.008	-0.007	-0.250	0.051
	(0.68)	(0.65)	(-0.53)	(-0.22)	(0.40)
SALES_VOL	0.033	0.038	-0.016	2.008	0.292
	(1.14)	(1.45)	(-0.68)	(1.08)	(1.43)
LNMV	-0.012***	-0.012***	-0.003	0.270	-0.036
	(-4.20)	(-4.34)	(-1.26)	(1.50)	(-1.42)
GM_RATIO				-5.543***	0.563***
				(-3.25)	(2.82)
FSALES				-0.887**	-0.148***
				(-2.42)	(-2.93)
CAPEX	-0.288*	-0.195	0.034		
	(-1.89)	(-1.30)	(0.25)		
Constant	0.206***	0.054	0.003	6.822***	0.191
	(5.60)	(1.51)	(0.09)	(4.57)	(0.72)
Observations	664	664	664	664	664
Adjusted R-squared	0.254	0.206	0.088	0.208	0.119
Year & Industry FE	YES	YES	YES	YES	YES

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Panel E Post-Pandemic Interactions

VARIABLES	(1) INVTO	(2) ABNINVTO	(3) RESINVTO	(4) INV_TURN	(5) INV_ELI
PAY_DISP	-0.003	-0.002	-0.003	-0.037	-0.018
	(-0.73)	(-0.57)	(-0.77)	(-0.24)	(-0.65)
CEO_DELTA	0.001	0.000	0.003	0.086	0.014
	(0.31)	(0.11)	(0.91)	(0.46)	(0.57)
CEO_VEGA	-0.001	-0.001	-0.002	-0.189	-0.001
	(-0.54)	(-0.63)	(-0.99)	(-1.27)	(-0.04)
BULLWHIP	0.007	0.011	0.024**	-0.986**	-0.220**
	(0.58)	(0.99)	(2.35)	(-2.33)	(-2.52)
PAY_DISP*H_BULLWHIP	-0.001	-0.001	-0.001	0.087***	0.012*
	(-1.46)	(-0.93)	(-0.66)	(2.60)	(1.68)
ROA	0.131***	0.106***	-0.049	-0.473	0.562**
	(3.58)	(3.02)	(-1.48)	(-0.31)	(2.11)
LEVER	-0.012	-0.026	-0.030	-0.004	0.077
	(-0.56)	(-1.24)	(-1.61)	(-0.00)	(0.46)
Q	0.001	0.001	-0.003*	0.057	0.032*
	(0.26)	(0.69)	(-1.68)	(0.50)	(1.81)
PPE	0.047***	0.030*	-0.005	1.152*	0.283**
	(2.61)	(1.77)	(-0.39)	(1.65)	(2.19)
SALES_GROW	0.009	0.008	-0.007	-0.245	0.052
	(0.66)	(0.64)	(-0.54)	(-0.22)	(0.41)
SALES_VOL	0.035	0.039	-0.015	1.869	0.273
	(1.21)	(1.49)	(-0.64)	(1.00)	(1.33)
LNMV	-0.012***	-0.012***	-0.003	0.265	-0.037
	(-4.19)	(-4.33)	(-1.25)	(1.47)	(-1.45)
GM_RATIO				-5.563***	0.561***
				(-3.26)	(2.81)
FSALES				-0.886**	-0.148***
				(-2.42)	(-2.94)
CAPEX	-0.274*	-0.187	0.040		
	(-1.80)	(-1.25)	(0.29)		
Constant	0.221***	0.062*	0.009	5.908***	0.063
	(5.71)	(1.67)	(0.25)	(3.64)	(0.22)
Observations	664	664	664	664	664
Adjusted R-squared	0.255	0.206	0.087	0.211	0.122
Year & Industry FE	YES	YES	YES	YES	YES

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

References:

Almutairi, A.R., Quttainah, M.A. and Almutairi, H.A., 2022. Inventory-management efficiency and consumer co-op financial performance. *Journal of Accounting, Finance & Management Strategy*, *17*(1), pp. 17-51.

Bannister, J.W., Newman, H.A. and Peng, E.Y., 2020. Top management tournament incentives and credit ratings. *Review of Quantitative Finance and Accounting*, *55*(2), pp.769-801.

Barker, J.M., Hofer, C., Hoberg, K. and Eroglu, C., 2022. Supplier inventory leanness and financial performance. *Journal of Operations Management*, *68*(4), pp.385-407.

Bendig, D., Strese, S. and Brettel, M., 2017. The link between operational leanness and credit ratings. *Journal of Operations Management*, *52*, pp.46-55.

Bonney, M.C., 1994. Trends in inventory management. *International Journal of Production Economics*, *35*(1-3), pp.107-114.

Callen, J.L., Fader, C. and Krinsky, I., 2000. Just-in-time: a cross-sectional plant analysis. *International Journal of Production Economics*, *63*(3), pp.277-301.

Chen, Y.F. and Disney, S.M., 2003. The order-up-to policy" sweet spot": using proportional controllers to eliminate the bullwhip problem. pp. 551-560.

Chen, H., Frank, M.Z. and Wu, O.Q., 2005. What actually happened to the inventories of American companies between 1981 and 2000?. *Management science*, 51(7), pp.1015-1031.

Chandra, C. and Grabis, J., 2005. Application of multi-steps forecasting for restraining the bullwhip effect and improving inventory performance under autoregressive demand. *European Journal of operational research*, *166*(2), pp.337-350.

Chuang, H. H. C., Oliva, R., & Heim, G. R. 2019. Examining the link between retailer inventory leanness and operational efficiency: Moderating roles of firm size and demand uncertainty. *Production and Operations Management*, 28(9), 2338–2364.

Core, J. and Guay, W., 2002. Estimating the value of employee stock option portfolios and their sensitivities to price and volatility. *Journal of Accounting research*, 40(3), pp.613-630.

Croson, R., Donohue, K., Katok, E. and Sterman, J., 2014. Order stability in supply chains: Coordination risk and the role of coordination stock. *Production and Operations Management*, *23*(2), pp.176-196.

Demeter, K. and Matyusz, Z., 2011. The impact of lean practices on inventory turnover. *International journal of production economics*, 133(1), pp.154-163.

Dolgui, A. and Ivanov, D., 2021. Ripple effect and supply chain disruption management: new trends and research directions. *International Journal of Production Research*, *59*(1), pp.102-109.

Eroglu, C. and Hofer, C., 2011. Lean, leaner, too lean? The inventory-performance link revisited. *Journal of Operations Management*, *29*(4), pp.356-369.

Feng, M., Li, C., McVay, S.E. and Skaife, H., 2015. Does ineffective internal control over financial reporting affect a firm's operations? Evidence from firms' inventory management. *The Accounting Review*, *90*(2), pp.529-557.

Green, J.R. and Stokey, N.L., 1983. A comparison of tournaments and contracts. *Journal of Political Economy*, *91*(3), pp.349-364.

Gupta, S., Starr, M.K., Farahani, R.Z. and Asgari, N., 2022. OM Forum—Pandemics/Epidemics: Challenges and Opportunities for Operations Management Research. *Manufacturing & Service Operations Management*, *24*(1), pp.1-23.

Hamm, S.J., Jung, B., Lee, W.J. and Yang, D.G., 2022. Organized Labor and Inventory Stockpiling. *The Accounting Review*, *97*(2), pp.241-266.

Kale, J.R., Reis, E. and Venkateswaran, A., 2009. Rank-order tournaments and incentive alignment: The effect on firm performance. *The Journal of Finance*, *64*(3), pp.1479-1512.

Kini, O. and Williams, R., 2012. Tournament incentives, firm risk, and corporate policies. *Journal of Financial Economics*, *103*(2), pp.350-376.

Ivanov, D., 2020. Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case. *Transportation Research Part E: Logistics and Transportation Review*, *136*, p.101922.

Ivanov, D. and Dolgui, A., 2021. OR-methods for coping with the ripple effect in supply chains during COVID-19 pandemic: Managerial insights and research implications. *International Journal of Production Economics*, *232*, p.107921.

Lazear, E.P. and Rosen, S., 1981. Rank-order tournaments as optimum labor contracts. *Journal of political Economy*, *89*(5), pp.841-864.

Lee, H.L., Padmanabhan, V. and Whang, S., 1997. The bullwhip effect in supply chains. *Sloan management review*, *38*, pp.93-102.

Lee, K.W., Lev, B. and Yeo, G.H.H., 2008. Executive pay dispersion, corporate governance, and firm performance. *Review of Quantitative Finance and Accounting*, *30*(3), pp.315-338.

Lin, Y., Liang, B. and Zhu, X., 2018. The effect of inventory performance on product quality: The mediating effect of financial performance. *International Journal of Quality & Reliability Management*. 35(10), pp. 2227-2247.

Mackelprang, A.W. and Malhotra, M.K., 2015. The impact of bullwhip on supply chains: Performance pathways, control mechanisms, and managerial levers. *Journal of Operations Management*, *36*, pp.15-32.

Mishra, S., Modi, S.B. and Animesh, A., 2013. The relationship between information technology capability, inventory efficiency, and shareholder wealth: A firm-level empirical analysis. *Journal of Operations Management*, *31*(6), pp.298-312.

Osadchiy, N., Schmidt, W. and Wu, J., 2021. The bullwhip effect in supply networks. *Management Science*, *67*(10), pp.6153-6173.

Petersen, M.A., 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of financial studies*, 22(1), pp.435-480.

Rozhkov, M., Ivanov, D., Blackhurst, J. and Nair, A., 2022. Adapting supply chain operations in anticipation of and during the COVID-19 pandemic. *Omega*, *110*, p.102635.

Shan, J., Yang, S., Yang, S. and Zhang, J., 2014. An empirical study of the bullwhip effect in China. *Production and Operations Management*, *23*(4), pp.537-551.

Shan, J. and Zhu, K., 2013. Inventory management in China: an empirical study. *Production and Operations Management*, *22*(2), pp.302-313.

Silver, E.A., 1981. Operations research in inventory management: A review and critique. *Operations Research*, *29*(4), pp.628-645.

Tsay, A.A. and Lovejoy, W.S., 1999. Quantity flexibility contracts and supply chain performance. *Manufacturing & Service Operations Management*, *1*(2), pp.89-111.

Wang, X. and Disney, S.M., 2016. The bullwhip effect: Progress, trends and directions. *European Journal of Operational Research*, *250*(3), pp.691-701.

Womack, J.P., Jones, D.T. and Roos, D., 2007. *The machine that changed the world: The story of lean production--Toyota's secret weapon in the global car wars that is now revolutionizing world industry*. Simon and Schuster.

Zighan, S., 2022. Managing the great bullwhip effects caused by COVID-19. *Journal of Global Operations and Strategic Sourcing*. 15(1), pp. 28047.

Marketing and Consumer Behavior

An examination of mobile payment adoption for Taiwanese consumers: A case of the retail industry

Yong-Cheng Chang

Ja-Shen Chen

Abstract

This study helps vendors and mobile payment service providers increase consumers' intention for mobile payment. The unified theory of acceptance and use of technology (UTAUT), technology, organization and environment (TOE) framework, trust, and perceived effectiveness of institutional e-commerce mechanisms (PEIEM) were integrated to examine consumers' intentions. The constructs and hypotheses were developed based on a literature review. Online questionnaires were used to collect data from Taiwanese consumers for mobile payment intentions. A total of 452 valid questionnaires were recovered for analysis on AMOS. The results showed that performance expectancy, effort expectancy, and social influence in UTAUT were positively correlated with online payment intention. Further, business organization and external environment, trust and PEIEM have positive impacts on consumers' intention. Trust and PEIEM also moderated the business organization and external environment and consumers' mobile payment intention. In sum, consumers are more likely to transact on mobile payment systems offered by trustworthy vendors with the process protected by laws and institutional e-commerce mechanisms against information concerns.

Keywords: mobile payment, unified theory of acceptance and use of technology (UTAUT), trust, perceived effectiveness of institutional e-commerce mechanisms (PEIEM)

I. Introduction

After the COVID-19 outbreak in 2020, consumers were concerned that the virus may spread through banknotes and coins. Meanwhile, lockdown or home quarantine measures imposed by countries worldwide have reduced face-to-face consumption activities, indirectly providing a big boost to online and mobile payment industries. With the growing ubiquity of smartphones, information technology has become an indispensable part of life for most people. Information system researchers have pursued a continuous quest to understand the factors influencing individuals' acceptance and use of emerging information technologies, and their efforts covered various theories, backgrounds, analyses, and methodologies (Dwivedi & Williams, 2008; Choudrie & Dwivedi, 2005; Williams et al., 2009). Venkatesh et al. (2003) developed the unified theory of acceptance and use of technology (UTAUT) in organizational settings. The discussion on technology acceptance models also covered cognitive, emotional, motivational, and behavioral intentions and individual responses (Momani, 2020; Hernandez 2017). The reasons why users accept or reject new technologies have become indispensable factors in the life cycle of any information system (Momani, 2020; Sivathanu and Pillai, 2019). Users' behavior was an emotional reaction to technology or information systems (Momani, 2020). Tornatzky and Fleischer (1990) also discussed the use of technology at the firm level and identified three factors that may affect the adoption of technology: technology, organization, and the environment (TOE). Technology refers to internal and external technologies relevant to a firm. Relative advantage, compatibility, complexity, trialability, and observability are considered technical factors that affect technology adoption (Rogers, 2014). The organizational context is associated with enterprise resources (organizational structures) (Oliveira and Martins, 2011). A firm's adoption of digital manufacturing or production may be affected by leadership support, organizational readiness, compatibility, and partners (Chatterjee et al., 2021). The adoption of artificial intelligence by organizations effectively increases productivity and helps people make quick decisions (Chatterjee et al., 2021, Duan et al., 2019; Dwivedi et al., 2021; Knight, 2015). Environment refers to the business environment in which a firm operates (industry, competitors, and technology service providers). Pressure from competitors, customers, and society are environmental factors that affect small and medium businesses' decisions about adopting new technologies (Ahmad et al., 2018; Low et al., 2011). Hellmueller and

Trilling (2012) also explained how trust mechanisms could be constructed from systems, procedures, or social transactions and how they affected people-to-people interdependencies. Assessing sources of information, media characteristics, and content indicated that trust does not exist without risk (Seligman, 1998). The perceived effectiveness of e-commerce institutional mechanisms (PEEIM) also illustrated that platform-based trust is built on trust in third-party structures (Pavlou & Gefen, 2004), independent of sellers and platforms. The sellers' specific mechanisms include website design and customer service (i.e., website quality), reducing social uncertainty and the psychological distance between online buyers and sellers. Based on a survey by the Institute for Information Industry in Taiwan, the rate of mobile payment usage in Taiwan was 62.2% in 2019, with an upward trend among the elderly. However, Taiwan has a low ratio of counterfeit currency and plenty of bank branches and automated teller machines. Users can use cash conveniently. E-tickets, financial cards, credit cards, and smartphones are widely used, and consumer behavior was affected by the pandemic.

II. Literature review and hypothesis development

1. Mobile payment in Taiwan

Electronic payment technology advances have lowered related equipment costs, allowing service providers to introduce multiple payment options for a convenient payment experience. In general, consumers and shopkeepers willingly accept electronic payments. Worldpay forecast that in 2022, over 50% of e-commerce and over 25% of physical sales worldwide will be through electronic wallets (E-Wallets) and similar means, with an average growth rate of 12%. Among other payment methods, there was a 2% increase in credit card usage and 1% increase in charge cards. Further, the proportion of cash payments decreased by 14%. Therefore, E-Wallet is expected to replace cash in small payments, while large payments will still be dominated by credit and charge cards. According to Dahlberg et al. (2008), Dahlberg, Guo and Ondrus (2015), Dermish et al. (2012), and Evans and Pirchio (2015), mobile payments refer to payments made by consumers with their mobile phones or devices through wireless networks and communication functions. They need to download a platform application and add a payment instrument (such as credit cards, financial cards, bank deposit accounts, or stored cash). Dahlberg and Mallat (2002) defined mobile payment as the payment for goods, services, and bills through mobile devices (such as mobile phones,

smartphones, and personal digital assistants) enabled by wireless technology or other communication technologies. Dahlberg (2008) suggested that mobile payment is payment for goods and services through communication technologies of wireless and other mobile devices. Kim et al. (2010) proposed that mobile payment is payment for goods, services, and bills through mobile devices. Based on the above definitions, it can be argued that mobile payment in Taiwan has become part of an integrated payment system. Businesses and consumers can conduct cashless transactions safely and conveniently with a single point-of-sale machine with wired or wireless payment systems, including accepting credit cards and a ticket vending machine.

2. UTAUT

Regarding the UTAUT proposed by Venkatesh et al. (2003), The discussion of cognitive, affective, motivational, and behavioral intentions and individual reactions (Momani, 2020; Hernandez 2017) suggests that the reasons why users accept or reject new technologies have become indispensable factors in the life cycle of any information system (Momani, 2020; Sivathanu and Pillai, 2019). The model identified possible factors influencing users' adoption, usage, and continued usage of new technologies (Momani, 2020). Factors such as usefulness, ease of use, complexity, and social influence have been identified for their effects on users' decisions about a technology and how it is used (Fishbein and Ajzen, 1975). This model explains users' intention to use an information system and usage behaviors. Four constructs influence behavioral intention and behaviors: (1) performance expectancy (PE), (2) effort expectancy (EE), (3) social influence (SI), and (4) facilitating conditions (FCs).

PE reflects the extent of perceived effective increase in work performance by the usage, comprising perceived usefulness, extrinsic motivation, job fit, relative advantage, and outcome expectations (Venkatesh et al., 2003). The payment system transfers money to shopkeepers for goods or services purchased, and mobile devices are used. Therefore, mobile payment involves various services, including account verification, initiation, scheduling payments, authentication, online payments, receipt of payments, and account management (Lai & Chuah, 2010). Therefore, in e-commerce, the payment system is an important component. Based on the above argument, it is evident that regarding the relationship between mobile payment and consumers, various mobile payment services provided by vendors may affect individuals' acceptance of mobile

payment. Hence, the following hypothesis was proposed.

Hypothesis 1: performance expectancy (*PE) positively and significantly affects users' adoption of mobile payment.*

Regarding EE, the ease of new technologies, systems, and operations for an individual refers to the simplicity of human-machine interfaces and operating systems (software), which affects the acceptance of technology by the individual (user) (Venkatesh et al., 2003). Furthermore, new technology brings safety and convenience to the relationship between vendors and consumers. For example, Kim et al. (2010) proposed mobile payments for goods, services, and bills through mobile devices in cashless transaction services. Based on the above argument, it is evident that for the relationships between mobile payment and consumers, mobile payment tools with easy operations and simple interfaces provided by vendors may affect individuals' acceptance. Therefore, the following hypothesis was proposed.

Hypothesis 2: Effort expectancy (EE) *positively and significantly affects users' mobile payment adoption.*

The SI construct measures others' effects on adopting and using new technologies. It comprises subjective norms, social factors, and public image. Subjective norms examine the effects of a particular image assigned to the individual, subject to the environmental states and mutual influences (Venkatesh et al., 2003). Social factors are subjective culture and collective consensus between individuals and other groups. A public image is an image of an individual that maintains or promotes one's status in the group—there is a significant positive correlation between the image and the norms endorsed by the reference group (Moore & Banbasat, 1991; Venkatesh et al., 2000). Based on the above argument, it is evident that individuals can be influenced to adopt mobile payment because of the image created, the subjective culture and coordinated consensus between the group and others, and the need to maintain or promote their status in the group. Therefore, the following hypothesis was proposed.

Hypothesis 3: social influence (SI) positively and significantly affects users' mobile

payment adoption.

The FCs construct measures the support organizations provide to individuals for new technology or new systems. The perceived behavioral control reflects the judgment about whether new technology or systems can be used from a purely behavioral perspective. FCs refer to providing technical assistance under a physical context. Compatibility is the degree to which new technologies are consistent with existing systems, values, or experiences (Venkatesh et al., 2003). Based on the above definitions, the relationship between mobile payment and consumers depends on whether the environment provided by the vendor for introducing new technologies for mobile payment is convenient. Additionally, relevant experiences may affect individuals' adoption of mobile payment. Therefore, the following hypothesis was proposed.

Hypothesis 4: Facilitating conditions (FCs) positively and significantly affects users' mobile payment adoption.

3. TOE

The TOE framework (shown in Fig. 1) proposed by Tornatzky and Fleischer (1990) has been applied to information technology research (Aboelmaged, 2014) and the analysis of e-commerce or e-commerce systems (Hong & Zhu, 2006). It can also be employed to explain factors in adoption decisions, including the explanation of modern technology adoption from the social, environmental, and technology context (Hossain & Quaddus, 2011). Any technology used in the social, environmental, and technology context can be easily explained using the TOE framework (Chatterjee et al., 2021, Hossain & Quaddus, 2011). Technology context deals with both internal and external effects (including benefits, characteristics of the system itself, innovativeness, and system complexity) of technologies on a firm. Both technologies have been used by the firm and are available on the market. Relative advantage, compatibility, complexity, trialability, and observability are considered technical factors that affect technology adoption (Rogers, 2014). Organizational context addresses organizational conditions (for example, firm size, organizational structure, the complexity of managerial hierarchy, and financial and human resources) (Oliveira and Martins, 2011). A firm's adoption of digital manufacturing or production may be affected by leadership support,

organizational readiness, organizational compatibility, and partners (Chatterjee et al., 2021). The adoption of artificial intelligence by organizations will effectively increase productivity and help people make quick decisions (Chatterjee et al., 2021, Duan et al., 2019; Dwivedi et al., 2021; Knight, 2015). Moreover, digitalization can make an organization smart (Chatterjee et al., 2021). Environmental context refers to industrial conditions and external stakeholders (competitors, customers, partners, governments, and the natural environment). Pressures from competitors, customers, and society are environmental factors that affect the adoption of new technologies by small and medium-sized businesses (Ahmad et al., 2018; Low et al., 2011). The biggest differences between private and public clouds are the bandwidth limitations imposed by suppliers and the improved efficiency of data management and processes to ensure the resilience and security of the cloud (Gangwar et al., 2015). Furthermore, some studies used TOE to verify whether information technology could achieve a high level of performance in organizational operations (Gangwar et al., 2015; Awa et al., 2017). The model was also used for managing data information systems (Giampietri & Trestini, 2020). In medical areas (Chang et al., 2007; Chong & Chan, 2012; Ahmadi, Nilashi, & Ibrahim, 2015), research has revealed the problems faced by medical institutions in adopting new technologies. Medical institutions must pay attention to the cost, safety, internal needs, organization size, and relative advantage (such as compatibility and complexity) of information adoption concerning patient personal information. Regarding organizations, they need to consider factors such as internal technical knowledge, financial resources, and the support of senior executives. The environment involves government policies, supplier support, competitive pressures, and market trends. These factors could affect medical institutions' adoption of new technologies or techniques. Based on the above argument, it is evident that technology, organization, and environment may affect individuals' adoption of mobile payment. Therefore, the following hypothesis was proposed.

Hypothesis 5: TOE positively and significantly affects users' adoption of mobile payment.

4. Trust

Technology adoption also depends on trust, which measures the extent to which

consumers are willing to accept vulnerability due to positive expectations of others' behavior or intentions (Hong & Cha, 2013). Kini and Choobineh (1998) studied trust in e-commerce settings. They found that all four dimensions of trust—information system, the task in the transaction process, information environment, and person—can influence trust. Lee and Turban (2001) revealed that consumers' trust in online shopping is influenced by whether vendors are trustworthy, and other factors. Qiu et al. (2003) documented that for the platform to be trustworthy, managers must provide good services and security features, including measures to safeguard consumers' personal information and services. Hence, the following hypothesis was proposed.

Hypothesis 6: Trust positively and significantly affects users' adoption of mobile payment.

According to Morgan and Hunt (1994), trust is a measure of the reliability and integrity of exchange partners. Hellmueller and Trilling (2012) also explained how trust mechanisms could be constructed from systems, procedures, or transactions and how they affected people-to-people interdependencies. The assessment of information sources, media characteristics, and content indicated that trust does not exist without risk (Seligman, 1998). Trust is a mental status (Kramer & Tyler, 1996). Based on the above argument, it is evident that trust in information sources and interpersonal dependency may affect individuals' adoption of mobile payment. We believe that the higher the level of trust show, the more TOE effect on users' adoption of mobile payment. Hence, the following hypothesis was proposed.

Hypothesis 7: Trust has a positively moderating effect between TOE and users' adoption of mobile payment.

5. PEEIM

In mobile payment transactions, PEEIM, the platform-based trust, is built on trust in third-party structures (Pavlou & Gefen, 2004), independent of sellers and platforms. The sellers' specific mechanisms include website design and customer service (i.e., website quality), which can reduce social uncertainty and the psychological distance between online buyers and sellers. Therefore, a more comprehensive understanding of

the moderating role of the institutional mechanisms of e-commerce and sellers' websites is required. They can mitigate the negative impact of risks and uncertainties in the online environment on transactions. Regarding suppliers and institutional levels, they are two distinct levels. Institutional mechanisms are generally independent on any particular online supplier or market environment at local levels for any particular online supplier. In an e-commerce environment, online customers are confident in the extent to which laws and technology protect them from privacy, identity loss, or credit card fraud through Internet protective measures. Based on the above argument, the level of Internet protection for mobile payment under institutional e-commerce mechanisms may affect individuals' adoption of mobile payment. Thus, the following hypotheses were proposed.

Hypothesis 8: PEEIM positively and significantly affects users' mobile payment adoption.

TOE also mentions the impact of leadership support, organizational readiness and partners in terms of organizational size, structure, management level, finance independent of any particular online supplier and local levels in a particular online supplier or market environment (Chatterjee et al., 2021 ; Rogers, 2014 ; Oliveira and Martins, 2011). TOE also mentions the impact of leadership support, organizational readiness and partners in terms of organizational size, structure, management level, finance and human resources, etc. Online customers trust the extent to which legal and technical protections on the Internet protect them from privacy or identity loss or credit card fraud. The higher the PEEIM protection mechanism of TOE, the higher the adoption of consumers' mobile payments. Based on the above argument, We believe that the higher the level of PEEIM show, the more TOE effect on users' adoption of mobile payment. Hence, the following hypothesis was proposed.

Hypothesis 9: PEEIM has a positive moderating effect between TOE and users' adoption of mobile.

Under such a context, this study examines how Taiwanese consumers' adoption of mobile payment is influenced from the perspectives of UTAUT, TOE, trust, and PEEIM

with the research framework shown in Fig. 1.

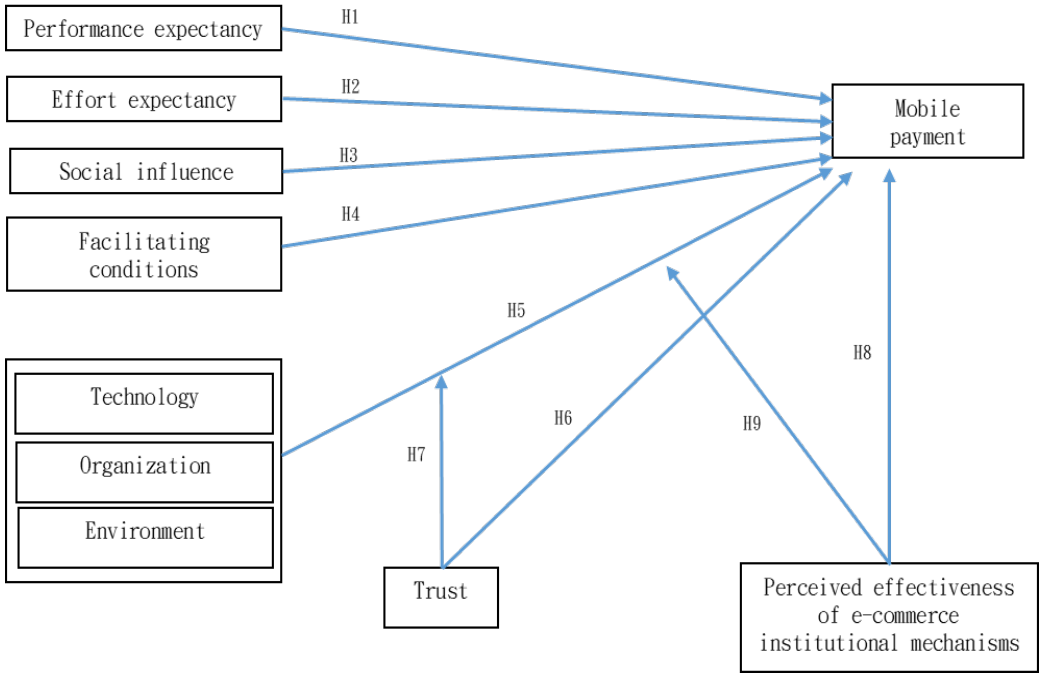


Figure 1. Research model

III. Methods

1. Sample and data collection

Data were collected through an online questionnaire hosted by SurveyCake. Eligible participants received a link to the questionnaire, and those who completed it were rewarded with LINE Points equivalent to a certain amount of cash. The questionnaire survey was conducted from April 8 to 18, 2022. A total of 485 questionnaires were recovered. Of these, 33 were considered invalid for the following reasons: the participant had no prior experience with mobile payment, and the data were incomplete or of poor quality (for example, all items were valued the same, notable difference between the location of consumption and the place of residence).

With 452 valid questionnaires recovered, the recovery rate was 93%. As shown by the demographics, 153 (33.8%) were males, and 299 (66.2%) were females. The largest age group comprised those aged 31 to 40 years (43.4%). Further, 73.9% had a college or university education, 41.4% worked in the service sector, and 45.4% earned a monthly income between NT$ 30,001 and NT$ 50,000. Moreover, 52% were employed in northern Taiwan, and 51.5% resided in northern Taiwan.

The mobile payment data showed that 74.6% of transactions were made using

credit cards added for mobile payment, and 30.0% were made through LINE Pay as a major payment system. Further, 27.7% of the mobile payment users used the service for 1–2 years, 31.0% for 2–3 years, and 44.2% daily. Moreover, 28.5% of the transactions were in the range NT$ 101–200 and 21.9% in NT$ 201–300. Additionally, 34.1% were for buying foods, and 51.5% of the transactions occurred in northern Taiwan.

2. Measures

All measurement items were adapted from previous studies, using the Likert five-point scale from 1 (strongly disagree) to 5 (strongly agree), with a higher score indicating a more agreeable statement (see Appendix for the statements). There were 16 statements for PE, EE, SI, and FC according to Venkatesh et al. (2003), nine for TOE by Tornatzky and Fleischer (1990) and Hsu, Ray, and Li Hsieh (2014), five for trust specified by Lewicki et al. (1998) and Lankton and McKnight (2011), and five for PEEIM according to Fang et al. (2014) to understand the operability of new technology for consumers, the trust in vendors or systems, effects of the information system on the organization and effects of changes in government decrees, and potential risks in electronic transactions. All items had been modified to adapt to the settings in which consumers adopt mobile payment.

IV. Data analysis and results

1. Measurement properties

Composite reliability (*CR*) and average variance extracted (*AVE*) were calculated to test the reliability and validity of the constructs. While *CR* is very similar to Cronbach's Alpha, it can measure the reliability of a scale more robustly (Netemeyer, Bearden & Sharma, 2003), and the value should be greater than 0.70 (Bagozzi & Yi, 1988). Further, the ideal *AVE* value should be greater than 0.50 (Hair et al., 1988). Except for TOE's *AVE* value of 0.46, less than 0.5 but acceptable, *CR* values for all constructs, as shown in Table 3, were well above the critical value, and the conditions for establishing convergence and discriminant validity were sufficiently satisfied.

Table 3. Validity and convergence validity

	CR	*AVE*	TOE	FC	SI	EE	PE	PIEEM	TRU	BI
TOE	0.884	0.461	0.679							
FC	0.846	0.581	0.799	0.762						
SI	0.839	0.568	0.511	0.525	0.754					
EE	0.889	0.615	0.740	0.795	0.408	0.784				
PE	0.887	0.611	0.764	0.764	0.395	0.847	0.782			
PIEEM	0.875	0.637	0.550	0.479	0.491	0.418	0.400	0.798		
TRUST	0.782	0.545	0.923	0.790	0.482	0.754	0.746	0.707	0.738	
BI	0.893	0.675	0.794	0.771	0.352	0.775	0.763	0.525	0.797	0.822

2. Model analysis

Regarding the fitness of a model, Schumacker and Lomax (2004) considered (x^2/df) less than five as a good fit. Gefen et al. (2000) and Hair et al. (1998) suggested that the goodness of fit index (*GFI*) and comparative fitness index (*CFI*) greater than 0.90 and the adjusted GFI (*AGFI*) greater than 0.8 indicate a good model fit. Jarvenpaa et al. (2000) and Browne and Cudeck (1993) indicated that the root mean square error of approximation (*RMSEA*) should ideally be less than 0.08 to suggest an acceptable fitness. The model fitness measures for this study are reported in Table 4.

Table 4 Measurement of model fitness

Fitness indexes	Ideal value	Recommended by	Value
Chi-square/degree of freedom (x^2/df)	< 5	Schumacker and Lomax (2004)	2.403
GFI	>0.8	MacCallum and Houng (1997)	0.853
AGFI	>0.8	Gefen et al. (2000) Hair et al. (1988)	0.829
CFI	>0.9	Gefen et al. (2000) Hair et al. (1988)	0.917
RMSEA	< 0.08	Jarvenpaa et al. (2000) MacCallum and Houng (1997)	0.056

3. Results for the direct effects

The model evaluation and confirmatory factor analysis were performed to determine the reliability, validity, and goodness of fit indexes. Path analysis was conducted to determine the correlation between all variables. Table 5 reports the

structural analysis results of the model on the paths from PE, EE, SI, FC, TOE, trust (TRUST), and PEEIM to mobile payment behavior intention (BI). The significant path coefficient (path coefficient = 0.268, t = 3.113, $p < 0.05$) from PE to BI supports hypothesis 1, and that (path coefficient = 0.264, t = 2.856, $p < 0.05$) from EE to BI supports hypothesis 2. Further, the significant path coefficient (path coefficient = 0.392, t = 4.739, $p < 0.01$) from SI to BI supports hypothesis 3, and that (path coefficient = 0.795, t = 13.415, $p < 0.001$) from TOE to BI supports hypothesis 5. Additionally, the significant path coefficient (path coefficient = 0.795, t = 12.946, $p < 0.001$) from TRUST to BI supports hypothesis 6, and that (path coefficient = 0.525, t = 9.777, $p < 0.001$) from PEEIM to BI supports hypothesis 7. Conversely, an insignificant path (path coefficient = -0.067, $p > 0.05$) from FC in UTAUT to BI does not support hypothesis 4. Figure 2 illustrates the path analysis results.

Table 5 Path analysis for constructs

		Standardized factor loading	Non-standardized factor loading	*SE.*	*CR.*	*p*
Performance expectancy	Behavior intention	0.268	0.279	0.09	3.113	**
Effort expectancy	Behavior intention	0.264	0.279	0.098	2.856	**
Social influence	Behavior intention	0.392	0.518	0.109	4.739	***
Facilitating conditions	Behavior intention	-0.067	-0.059	0.038	-1.542	0.123
Technology, organization, and environment	Behavior intention	0.795	0.964	0.072	13.415	***
Trust	Behavior intention	0.780	0.976	0.075	12.946	***
Perceived effectiveness of e-commerce institutional mechanisms	Behavior intention	0.525	0.553	0.057	9.777	***

Note: SE.(Standard Error), CR.(Critical Ratio）

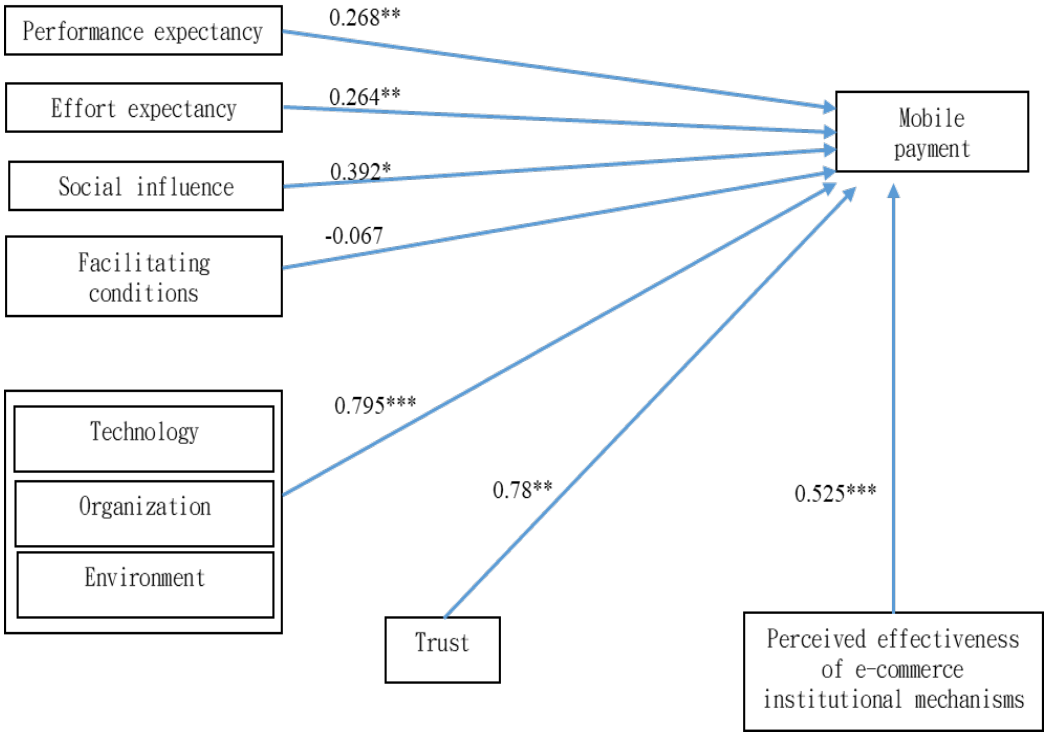


Figure 2. Path analysis

4. Moderating effects of trust and PEEIM

TOE was used to evaluate the moderating effects of trust on BI. The interaction terms (TOE × TRUST, TOE × PEEIM) are the product of independent variables and moderators, and the mean of these results was tested by AMOS 22. The interaction term TOE × TRUST has a positive and significant moderating effect on the relationship of BI (path coefficient = 0.534, t = 5.641, p < 0.001), supporting hypothesis 7. The interaction term TOE × PEEIM has a positive and significant moderating effect on the relationship of BI (path coefficient = 0.686, t = 7.403, p < 0.001), supporting hypothesis 9. The detailed results are presented in Table 6.

Table 6. Moderating effects of TRUST and PEEIM

		Standardized factor loading	Non-Standardized factor loading	*SE.*	*CR.*	*p*
Trust	Behavior intention	0.534	0.078	0.014	5.641	***
Perceived effectiveness of e-commerce institutional mechanisms	Behavior intention	0.686	0.051	0.007	7.403	***

Note: SE.(Standard Error), CR.(Critical Ratio）

V. Discussion and conclusions

This study shows that PE positively and significantly affects users' adoption of mobile payment (hypothesis 1). Users reported that vendors' mobile payment systems increased checkout speed, consistent with the needs of daily life and the prevailing business practices. EE positively and significantly affects users' mobile payment adoption (hypothesis 2). Regarding the mobile payment system, users indicated that simple interfaces, ease of operation, and intuitive interactions were important and could affect users' adoption of the technology. SI positively and significantly affects users' mobile payment adoption (hypothesis 3). Others can influence users and are concerned about assistance from mobile payment vendors in addressing system issues. Better assistance is conducive to users' adoption of mobile payment. FC has no significant impact on users' mobile payment adoption (hypothesis 4). The finding indicates that users believe vendors do not provide a favorable environment for mobile payment, the operation is not facilitated, and the user experience is not satisfactory. While vendors believe that they have provided a favorable environment and good facilities, the belief is not shared by users while using mobile payment. Moreover, other factors, such as different user characteristics and living environments, may also hinder the positive relationship between FC and users' mobile payment adoption. TOE positively and significantly affects users' mobile payment adoption (hypothesis 5). Users are concerned with mobile payment technology (system simplification and innovation), organization (size and financial conditions), environment (industrial status, government regulations, and partners) and pressure from competitors, customers, and society. These factors affect their adoption of mobile payment. TRUST positively and significantly affects users' mobile payment adoption (hypothesis 6). Users believe trusting information sources and interpersonal dependency can affect mobile payment adoption. Furthermore, with TRUST, TOE positively affects users' mobile payment adoption (hypothesis 7). If users trust that the mobile payment system is technologically effective in facilitating vendors using new and innovative business models to improve service quality and streamline transaction flow, it can reduce human errors in the transaction process, providing customers with a new experience. Additionally, in terms of environment, government regulations and a secure environment can provide mechanisms for protecting transaction data, and the rate of mobile payments can be increased effectively. This suggests that TRUST moderates the relationship between

TOE and BI. In other words, with a higher level of trust, users have higher rates of using mobile payments. PEEIM positively and significantly affects users' mobile payment adoption (hypothesis 8). Users value the platform transaction mechanisms, and these platforms must ensure that laws and technologies protect users by guarding against privacy or credit card fraud issues. With PEEIM, TOE positively affects users' adoption of mobile payment (hypothesis 9). If users believe that the transactional platform can effectively maintain user information security (system simplification and innovation), organization (size and financial conditions), and environment (industrial status, government regulations, and partners), users will use mobile payment more. This suggests that PEEIM moderates the relationship between TOE and BI. In other words, with a higher level of TOE due to PEEIM, users are more likely to trade through the platform.

1. Theoretical implications

This study is significant as it analyses consumer adoption of mobile payment from four perspectives. First, it integrated UTAUT, TOE, TRUST, and PEEIM to test hypotheses about consumers' mobile payment adoption. Although previous researchers have proven the importance of information systems or transactional processes, few studies have verified the effects of trust and PEEIM in mobile payment settings. Therefore, a questionnaire was developed to examine mobile payment adoption by Taiwanese consumers. Second, trust reflects the willingness of consumers to use mobile payment. The adoption rate can be affected by whether shopkeepers have provided trustworthy transaction systems. The analysis shows that vendors can enhance service quality and bring new service experiences while protecting users' transaction data by providing secure transaction processes. Third, this study verified that users pay attention to the data security of the platform transaction process regarding institutional e-commerce mechanisms. Therefore, e-commerce mechanisms should conform to legal requirements and have security features to ensure safety and privacy in transactions. This move can increase mutual trust between vendors and consumers, continuously attracting users to adopt the technology. Fourth, the empirical results of the study show that trust and institutional e-commerce mechanisms independently moderate the relationship between TOE and BI. Regarding transaction processes, users care about improved service quality, reduced transaction errors, and information protection. Few

studies have examined the moderating effects of trust and institutional e-commerce mechanisms on the relationship between TOE and BI. Nevertheless, the insights from this study can extend the literature on consumers' adoption of mobile payment. Therefore, the effects of trust and e-commerce institutional mechanisms, as quantified in this study, make a special contribution to the literature.

2. Managerial implications

This study provides valuable information to vendors and mobile payment developers. Through a comprehensive framework of UTAUT, TOE, trust, and PEEIM, the study can improve the understanding of consumers' BI and issues of interest. The analysis results can inform vendors and mobile payment developers regarding future strategies to enhance mobile payment systems for consumption. First, users adopting mobile payments across UTAUT factors (PE, EE, SI, and FC) should use a mobile payment system that increases speed, has simple interfaces, an easy-to-operate and intuitive system, and a vendor guarantee in fixing system issues. These are conducive to consumers' mobile payment adoption. Second, regarding TOE, users are concerned about whether the technology for mobile payment is simple and easy to use, the organization and the scale of mobile payment operation, the industrial status, and whether the government law protects users, shopkeepers, and developers. Third, vendors and stakeholders should provide users with a system they can trust for mobile payment. Interpersonal relationships determine users' trust. When vendors and mobile payment stakeholders provide users with a trustworthy system, BI will increase. Fourth, regarding PEEIM, the findings suggest that effective protection of user data can increase the adoption rate. Finally, users are concerned about whether laws and network technology protect them against privacy or credit card issues during transactions.

3. Limitations and future research

First, this study examined Taiwanese consumers' BI for adopting mobile payment. Second, although it uses quantitative methods for analytical research, future studies may use quantitative methods and in-depth interviews to draw more descriptive information from the respondents. Third, a statistical analysis of the collected data shows that *AVE* for TOE is lower than the threshold value but is still acceptable. The validity concern may be addressed in future research. Additionally, in terms of the

organization, future research may focus on whether mobile payment systems can reduce the required number of staff, avoid transaction errors, and bring new service experiences for users (consumers).

References

Aboelmaged, M. G. (2014). Predicting e-readiness at firm-level: An analysis of technological, organizational and environmental (TOE) effects on e-maintenance readiness in manufacturing firms. *International Journal of Information Management*, 34(5), 639-651.

Ahmad, S. Z., Bakar, A. R. A., & Ahmad, N. (2018). Social media adoption and its impact on firm performance: the case of the UAE. *International Journal of Entrepreneurial Behavior & Research.*

Ahmadi, H., Nilashi, M., & Ibrahim, O. (2015). "Organizational decision to adopt hospital information system: An empirical investigation in the case of Malaysian public hospitals." *International journal of medical informatics*, 84(3), 166-188.

Ajzen, I., & Fishbein, M. (1975). A Bayesian analysis of attribution processes. *Psychological bulletin*, 82(2), 261.

Awa, H. O., Ojiabo, O. U., & Orokor, L. E. (2017). Integrated technology-organization-environment (TOE) taxonomies for technology adoption. *Journal of Enterprise Information Management*.

Bag, S., Pretorius, J. H. C., Gupta, S., & Dwivedi, Y. K. (2021). Role of institutional pressures and resources in the adoption of big data analytics powered artificial intelligence, sustainable manufacturing practices and circular economy capabilities. *Technological Forecasting and Social Change*, 163, 120420.

Bagozzi, R. P.; Yi, Y. On the evaluation of structural equation models. *Journal of the Academy of Marketing Science* 1988, 16, 74-94.

Chang, I. C., Hwang, H. G., Hung, M. C., Lin, M. H., & Yen, D. C. (2007). "Factors affecting the adoption of electronic signature: Executives' perspective of hospital information department." *Decision support systems*, 44(1), 350-359.

Chatterjee, S., Rana, N. P., Dwivedi, Y. K., & Baabdullah, A. M. (2021). Understanding AI adoption in manufacturing and production firms using an integrated TAM-TOE model. *Technological Forecasting and Social Change*, 170, 120880.

Chong, A. Y.-L., & Chan, F. T. S. (2012). "Structural equation modeling for multi-stage analysis on radio frequency identification (RFID) diffusion in the health care industry". *Expert Systems with Applications*, 39(10), 8645-8654. doi:10.1016/j.eswa.2012.01.201

Choudrie, J., & Dwivedi, Y. K. (2005). Investigating the research approaches for examining technology adoption issues. *Journal of Research Practice,* 1(1), D1-D1.

Dahlberg, T., & Mallat, N. (2002). Mobile payment service development-managerial implications of consumer value perceptions. ECIS 2002 Proceedings, 139.

Dahlberg, T., Guo, J., & Ondrus, J. (2015). A critical review of mobile payment research. *Electronic Commerce Research and Applications*, 14(5), 265-284.

Dahlberg, T., Mallat, N., Ondrus, J., & Zmijewska, A. (2008). Past, present and future of mobile payments research: A literature review. *Electronic commerce research and applications*, 7(2), 165-181.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management science*, 35(8), 982-1003.

Dermish, A., Kneiding, C., Leishman, P., & Mas, I. (2012). Branchless and Mobile Banking Solutions for the Poor: A Survey of Literature, *Innovations*, Vol. 6(4), 81-98.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda. *International journal of information management*, 48, 63-71.

Dwivedi, Y. K., & Williams, M. D. (2008). Demographic influence on UK citizens'e-government adoption. *Electronic Government, an International Journal*, 5(3), 261-274.

Evans, D., & Pirchio, A. (2015). An empirical examination of why mobile money schemes ignite in some developing countries but flounder in most. Coase-Sandor Institute for Law and Economics Working Paper, 723. The University of Chicago Law School.

Fang, Y., Qureshi, I., Sun, H., McCole, P., Ramsey, E., & Lim, K. H. (2014). Trust, satisfaction, and online repurchase intention. *Mis Quarterly*, 38(2), 407-A9.

Gangwar, H., Date, H., & Ramaswamy, R. (2015). Understanding determinants of cloud computing adoption using an integrated TAM-TOE model. *Journal of enterprise information management*.

Giampietri, E., & Trestini, S. (2020). Analysing farmers' intention to adopt web marketing under a technology-organisation-environment perspective: A case study in Italy. Agricultural Economics, 66(5), 226-233.

Hair, J. F.; Anderson, R. E.; Tatham, R. L.; William, C. Black (1998), Multivariate Data Analysis. In: Upper Saddle River, NJ: Prentice Hall, 1998.

Hellmueller, L. and Trilling, D. (2012). The credibility of credibility measures: A meta-analysis in leading communication journals, 1951 to 2011. In WAPOR 65th Annual Conference in Hong Kong.

Hernandez, A. A. (2017). Green IT Adoption Practices in Education Sector: A Developing Country Perspective. *International Journal of Sociotechnology and Knowledge Development (IJSKD)*, 9(3), 1-16.

Hong, I. B., & Cha, H. S. (2013). The mediating role of consumer trust in an online merchant in predicting purchase intention. *International Journal of Information Management*, 33(6), 927-939.

Hong, W., & Zhu, K. (2006). Migrating to internet-based e-commerce: Factors affecting e-commerce adoption and migration at the firm level. *Information & management*, 43(2), 204-221.

Hossain, M. A., & Quaddus, M. (2011). The adoption and continued usage intention of RFID: an integrated framework. *Information Technology & People*.

Kim, C., Mirusmonov, M., & Lee, I. (2010). An empirical examination of factors influencing the intention to use mobile payment. *Computers in Human Behavior,* 26(3), 310-322.

Kini, A., & Choobineh, J. (1998, January). Trust in electronic commerce: definition and theoretical considerations. In Proceedings of the thirty-first Hawaii International conference on System sciences (Vol. 4, pp. 51-61). *IEEE*.

Knight, R. (2015). Convincing skeptical employees to adopt new technology. *Harvard Business Review*, 2, 1-7.

Lai, P. M., & Chuah, K. B. (2010, October). Developing an analytical framework for mobile payments adoption in retailing: a supply-side perspective. In 2010 International Conference on Management of e-Commerce and e-Government (pp. 356-361). *IEEE*.

Lee, M. K., & Turban, E. (2001). A trust model for consumer internet shopping. *International Journal of electronic commerce*, 6(1), 75-91.

Low, C., Chen, Y., & Wu, M. (2011). Understanding the determinants of cloud computing adoption. *Industrial management & data systems.*

Momani, A. M. (2020). The unified theory of acceptance and use of technology: A new approach in technology acceptance. *International Journal of Sociotechnology and Knowledge Development (IJSKD)*, 12(3), 79-98.

Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information systems research*, 2(3), 192-222.

Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of marketing*, 58(3), 20-38.

Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). Scaling procedures: Issues and applications. sage publications.

O'Hair, D., Friedrich, G. W., & Dixon, L. D. (1998). Strategic communication in business and the professions.

Oliveira, T., & Martins, M. F. (2011). Literature review of information technology adoption models at firm level. *Electronic Journal of Information Systems Evaluation*, 14(1), pp110-121.

Pavlou, P. A., & Gefen, D. (2004). Building effective online marketplaces with institution-based trust. *Information systems research*, 15(1), 37-59.

Qiu, L., Yang, Y. R., Zhang, Y., & Shenker, S. (2003, August). On selfish routing in Internet-like environments. *In Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications* (pp. 151-162).

Rogers, E. M., Singhal, A., & Quinlan, M. M. (2014). Diffusion of innovations. In An integrated approach to communication theory and research (pp. 432-448). Routledge.

Schumacker, R. E., & Lomax, R. G. (2004). A beginner's guide to structural equation modeling. psychology press.

Seligman, M. E. (1998). Building human strength: Psychology's forgotten mission.

Sivathanu, B., & Pillai, R. (2019). Leveraging technology for talent management: Foresight for organizational performance. *International Journal of Sociotechnology and Knowledge Development (IJSKD)*, 11(2), 16-30.

Tornatzky, L. G., Fleischer, M., & Chakrabarti, A. K. (1990). Processes of technological innovation. Lexington books.

Tyler, T. R., & Kramer, R. M. (1996). Whither trust. Trust in organizations: Frontiers of theory and research, 1, 15.

Venkatesh, V., Morris, M. G., & Ackerman, P. L. (2000). A longitudinal field investigation of gender differences in individual technology adoption decision-making processes. *Organizational behavior and human decision processes*, 83(1).

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

Williams, L. J., Vandenberg, R. J., & Edwards, J. R. (2009). 12 structural equation modeling in management research: A guide for improved analysis. Academy of Management Annals, 3(1), 543-604.

National Development Council [online] https://www.facebook.com/ndc.gov.tw/photos/a.1522479164686786/1969042433363788/?type=3/ (Accessed 21 August 2021).

Would Gen Z Enjoy VR Virtual Tourism Experience? The Effects of Virtual Tourism Characteristics on Behavioral Intention

Abstract

Purpose- VR virtual tourism has become popular due to the impact of the COVID-19 epidemic and has become an essential alternative to physical tourism. There is a lack of knowledge about the relationship between VR virtual tourism characteristics, customer experiential value, and Gen Z's behavioral intention.

Design/methodology/approach-This study collected 409 valid online questionnaires through social media platforms and analyzed data by using AMOS 27 to understand the extrinsic and intrinsic value of the consumer experience.

Findings- The results show that the characteristics of VR virtual tourism (i.e., imagination, immersion, content quality cognition, and perceived quality) have a positive influence on customer experiential value. In addition, customer experiential value has a positive influence on behavioral intention.

Originality/value-Based on the 3I characteristics of VR design, this study validates the construct of VR virtual tourism. In addition, this study proposes a conceptual framework as the "VR experiential value model" and contributes to theoretical development while providing substantive recommendations for tourism service companies to attract Gen Z.

Keywords: VR, virtual tourism, customer experiential value, behavioral intention, Gen Z

Introduction

With the advancement and popularization of AI and AR/VR technologies, the industry has used cloud computing, big data, 5G, and other technologies to provide consumer services, and the transition between online and offline has become increasingly smooth. The immersive interaction of an increasing number of Gen Z in pursuit of innovation has become an essential topic of service innovation. Gen Z, born between 2000 and 2010, are called digital natives. They are increasingly interested in virtual interactive experiences and become increasingly interested in games, movies, and amusement parks. As a result, many VR wearable device innovations (such as Oculus Quest, HTC Vive, and SONY PlayStation VR) and mixed reality (MR) interfaces (such as HoloLens) have led to major technological innovation growth, giving Gen Z more participation and diversified choices (Buhalis et al., 2019).

The 2021 epidemic not only affects the development, process, and sustainability of the tourism industry but also affects tourists' travel attitudes, travel risk perception, intention, and behavior (Ntounis et al., 2022). This postepidemic era has led to market changes and the emergence of alternative models (Ioannides and Gyimóthy, 2020; Kwok and Koh, 2020), such as the use of VR cloud tourism. Therefore, VR technology has become an important platform for tourism enterprises to maintain revenue sources and promote them (Itani and Hollebeek, 2021). Although virtual tourism using the Internet and 360 videos already exists, due to the epidemic situation and the progress of 5G streaming technology, VR has recently become a popular marketing tool in the tourism industry. VR technologies can be classified according to immersion, presence, and complexity (Beck and Egger, 2018). Immersive classification can be defined as nonimmersive (e.g., Imax movie), semiimmersive (e.g., 360 projection), and fully immersive (e.g., VR glasses, head-mounted display) based on the type of environment simulation and the user's perception of the real world (Bafadhal and Hendrawan, 2019). This study focuses on a completely immersive VR virtual tourism environment, which can provide a sense of immersion and telepresence by using digital images of actual tourist attractions or an unreal but theoretical space (Huang et al., 2016). It is

conveyed to virtual tourists (also known as remote tourists) so that users feel that they are immersive and real (Stepaniuk et al., 2015); users can immerse themselves, look around and control the experience (Yung and Khoo-Lattimore, 2019).

Previous studies on virtual tourism are mostly based on the universal development of perceived ease of use and perceived usefulness in the technology acceptance model (TAM) (Schiopu et al., 2021) or are mainly based on qualitative research (Buhalis and Karatay, 2022). The method may be subjective because these small sample-based conclusions cannot provide broad explanatory power. In addition, there is still a huge difference between the characteristics of VR virtual tourism and real tourism. When Gen Z is unable to travel abroad, what is their true behavioral intention toward VR virtual tourism? It is still difficult to determine a clear response.

When it is impossible to participate in real tourism, VR virtual tourism is an effective choice. It may be possible for Gen Z to learn the original history of scenic spots, how to protect cultural heritage, and how to provide adventurous tourism services for tourists with limited mobility, financial difficulties, or time constraints (Yung and Khoo-Lattimore, 2019). However, VR virtual tourism cannot obtain high-quality online five-senses experience because all sense organs must be in contact with the physical environment to be considered a real experience (Kim et al., 2020); in addition, only a few scholars have studied the epidemic situation. Therefore, there is still a lack of research on the cognition and behavior of Gen Z regarding the use of VR virtual tourism.

This paper aims to contribute to the current debate on the antecedents of customer experiential value in VR virtual tourism content and on the effect of customer experiential value on behavioral intention (Figure 1). In doing so, our research focuses on three specific types of VR characteristics in virtual tourism: immersion, imagination, and interactivity. There is a lack of empirical studies that provide evidence of the relationships between these VR characteristics and experiential customer value, particularly in the context of VR virtual tourism, and examine the ultimate effect of experiential customer value on behavioral intention.

Literature review and hypothesis development

VR virtual tourism characteristics

Successful VR design requires three characteristics (Burdea and Coiffet, 2003): immersion, imagination, and interactivity. *Immersion* refers to the sensory device worn by the user in VR to generate visual and auditory stimuli to achieve immersive and believable perception (Sherman and Craig, 2018). *Imagination* refers to using personal images in a virtual environment to form an imaginary space that is different from the natural environment (Lee and Jeong, 2021). *Interactivity* refers to the user's response to the virtual environment content through the operation interface of the sensing device, and the VR system provides the reaction based on the user's response.

In terms of interactivity, this study proposes that there are two major dimensions of content quality cognition and perceived quality of the trip. *Content quality cognition*: The cognition here has two meanings: first, the cognitive value of different tour contents to the same customer is comparable; second, the customer value of the same tour content varies from person to person because even for the same tour content, the customer experience is also different (Carter and Gilovich, 2010). *Perceived quality*: customer value represents the customer's perceived interactive quality. The intrinsic aspect emphasizes the hedonic consumption of hedonic and playfulness rather than the result (Wei et al., 2016), attitude (like and dislike), evaluation (good and bad), opinion (for and against), reaction tendency (approaching and avoiding), or value (positive and negative). Therefore, a real-time response is needed to achieve human-computer interaction and cognition. Customer value is generated in the interaction between the customer and the product rather than created in advance by the product supplier (Rahimizhian et al., 2020).

Imagination (IMA)		**Sources**
IMA1	The VR virtual tour provided by the tour content is comfortable	Burdea and Coiffet (2003)
IMA2	The VR virtual tour content provided by this tour content is overall smooth.	
IMA3	After VR virtual tourism, there will be more imagination space for the destination.	

IMA4	After the VR virtual tour, the product content of the VR virtual tour has a sense of novelty.	
Immersion (IMM)		
IMM1	The immersive ability of the VR virtual tour provided by the tour content can make the customers satisfied.	Burdea and Coiffet (2003)
IMM2	The immersive sense of the VR virtual tour provided by this tour content will bring customers an immersive experience.	
IMM3	The immersiveness of the VR virtual tour provided by the tour content will allow customers to focus more on the experience.	
IMM4	The immersion provided by VR virtual tourism will make you completely immersed in the virtual world and cannot distinguish between reality and the virtual world.	
Content quality cognition (CQC)		
CQC1	Will choose the interactive content experience of this tour because of the relationship of tourism.	Verhoef et al. (2009)
CQC2	Participated in the course content and induced emotional connection.	
CQC3	After experiencing the VR interactive experience provided by the tour content, the impression of the travel agency was indeed enhanced.	
CQC4	Have participated in the interactive experience activities of the itinerary and purchased the itinerary.	
Perceived quality (PQ)		
PQ1	The interactive experience provided by the content of the tour is high	Aaker (1991); Petrick (2002)
PQ2	The interactive experience provided by the content of the tour is trustworthy.	
PQ3	The information in the interactive experience provided by the tour content is reliable.	
PQ4	The interactive experience and service provided by the content of the tour will want to purchase real travel products.	

Consumer return on investment (CROI)		
CROI1	The VR virtual tour provided by the tour content allows customers to get their due value after payment and time.	Mathwick et al. (2001); Parasuraman and Grewal (2000)
CROI2	The VR virtual tour provided by the tour content will make customers want to make subsequent purchases after paying time.	
CROI3	The VR virtual tour provided by the tour content is economical for the customer.	
CROI4	Under the consideration of price and time, it is considered that the selection of the course content is the right decision.	
Aesthetics (AE)		
The VR virtual tour provided by the tour content……		
AE1	can give customers a brand-new sensory experience.	Mathwick et al. (2001); Pullman and Gross (2004); Wang et al. (2011)
AE2	always provides value-added services and virtual content.	
AE3	will make customers want to use it again.	
AE4	will make customers feel comfortable.	
Service excellence (SE)		
SE1	The VR virtual tour and service provided by the tour content are of consistent quality.	Mathwick et al. (2001); Sweeney and Soutar (2001)
SE2	The VR virtual tour provided by the tour content can enhance the customer's sense of identity with them.	
SE3	The price of the VR virtual tour provided by the tour content is reasonable.	
SE4	Customers will continue to support and purchase if the content of the tour will launch other experience content in the future.	
Playfulness (PF)		
PF1	The VR virtual tour provided by this tour content is to make customers feel happy and enjoyable.	Mathwick et al. (2001); Overby and Lee (2006); Bilgihan et al. (2016)
PF2	The VR virtual tour provided by this tour content will make customers forget the temporary annoyances.	
PF3	The content of the tour provides VR virtual tour and the service efficiency is in line with the original expectation of customers.	
PF4	The VR virtual tour provided by this tour content makes customers want to become the value cocreator of virtual content.	
Behavioral intention (BI)		
BI1	If there is a chance, we will recommend the tour content provided by the travel agency to other people.	Zeithaml et al. (1996); Bhattacherjee, (2004); Moon and Kim (2001)
BI2	If you have the opportunity, you will come to experience the VR virtual tour regularly.	
BI3	If you have the opportunity, you will recommend others to experience the VR virtual tour.	
BI4	When someone seeks advice, they will be recommended to experience the VR virtual tour.	
BI5	Will encourage friends or family members to experience VR virtual tourism	

Customer experiential value

Holbrook (1994) proposed four dimensions of customer experience, namely, consumer return on investment, service excellence, aesthetics, and playfulness. *Consumer return on investment* refers to the rate of return on customer investment, and the financial, behavioral, and psychological resources that can provide returns are called customer return on investment (Mathwick et al., 2001). *Service excellence* refers to the responsiveness of the service to meet customer needs. *Aesthetic* is the response to the performance of service work, the control of the object, and the balance ratio (Olson, 1972). Aesthetics have two directions: entertainment and dramatic service performance and satisfaction with visual environmental elements (Holbrook, 1982). *Playfulness* is generated internally by customers and is self-oriented; it is the happiness obtained using products or services (Chou and Lu, 2009).

VR technology creates a sense of virtual presence, which is a psychological state that exists in the intermediary environment. The virtual presentation of images will stimulate people to actively participate in mental image processing (Coyle and Thorson, 2001; Klein, 2003; Suh and Chang, 2006; Tussyadiah et al., 2018). The concept of mental imagery is defined as "a process of imagination through which sensory information is presented in memory" (MacInnis and Price, 1987). When people imagine something, they will form a psychological image based on previous experience and/or available information; tourism research records the role of imagination as the core driving force for positive customer outcomes (Lee and Gretzel, 2012). Therefore, this study proposes the following hypothesis: *H1: The imagination characteristic of VR virtual tourism positively influences customer experience value.*

For almost all tourism products and services, the choices are driven by the expected future hedonic benefits, such as happiness, relaxation, or excitement (Kwon and Lee, 2020). If the vivid performance in VR is similar to the experience in real life, customers can use them to more easily imagine the experience of specific objects or settings displayed in VR, thereby more in-depth prediction of the feeling that these objects/settings will bring to them. Because emotion prediction

involves the creation of mental representations, factors that affect this creative process may also affect emotion prediction (Park and MacInnis, 2006). Previous studies have suggested that advertisements can evoke psychological images, thereby positively affecting customers' responses to products (Yoo and Kim, 2014). Therefore, this study proposes the following hypothesis: *H2: The immersion characteristic of VR virtual tourism positively influences customer experience value.*

The interaction with the experience content can also stimulate more stimulating emotions among customers (Persky and Blascovic, 2007). The cognition of the tour content refers to whether the cognition and idea of the original tour are different after the VR virtual tour experience, and then the amount of information received during the experience is changed. During the interactive virtual tour experience, if the customer has a bad first impression of the tour or does not understand the tour, he or she will be more willing to receive more information when learning about the tour or product through the interactive features. If the customer already knows a lot about the tour or product, he or she will be more willing to purchase the product after the interactive experience of VR, thereby generating the intention to use the product (Grace and O'Cass, 2004; Verhoef et al., 2009; Lemke et al., 2011). Therefore, this study proposes the following hypothesis: *H3: The content quality cognition of VR virtual tourism positively influences customer experience value.*

Perceived quality refers to whether the customer associates with others during the interactive experience and then affects the purchase decision after the experience, and the expectation before the trip affects a person's travel decision (Del Bosque et al., 2006). This study predicts that the behavioral impact of using VR virtual tourism will be affected by the customer's previous experience of the travel destination. Experienced customers may have an existing psychological impression of the destination based on their direct experience, and the attitude that generates direct experience is more confident than the attitude formed by indirect experience (Rajagopal and Montgomery, 2011). In addition, when experienced customers watch VR content, they may not pay much attention to the characteristics of the attractions (Buehler and McFarland, 2001), so the mental imagery induced by VR will be covered by previous experience. Therefore, compared with

experienced customers, inexperienced customers may be more strongly affected by VR interactivity (Garvin, 1983; Steenkamp, 1990; Aaker, 1991). Therefore, this study proposes the following hypothesis: *H4: The perceived quality of VR virtual tourism positively influences customer experiential value.*

Taking the tourism industry as an example, after experiencing VR virtual tourism, customers will finally intend to use the itinerary or tickets because of the above four points. Li Jinwei, the founder of Trip Moment (2020) has motioned" *Travel is a "low-frequency, high-consumption product, but there is no way to try it."* This is the pain point of the travel industry; early use of VR layout can help the industry establish a sense of trust in the minds of customers. VR can be a substitute for a real journey: "It is an entertainment experience of travel, and VR travel is provided to customers as a low-consumption, high-frequency entertainment." Therefore, this study infers the following hypothesis: *H5: Customer experience value has a positive impact on behavioral intention.*

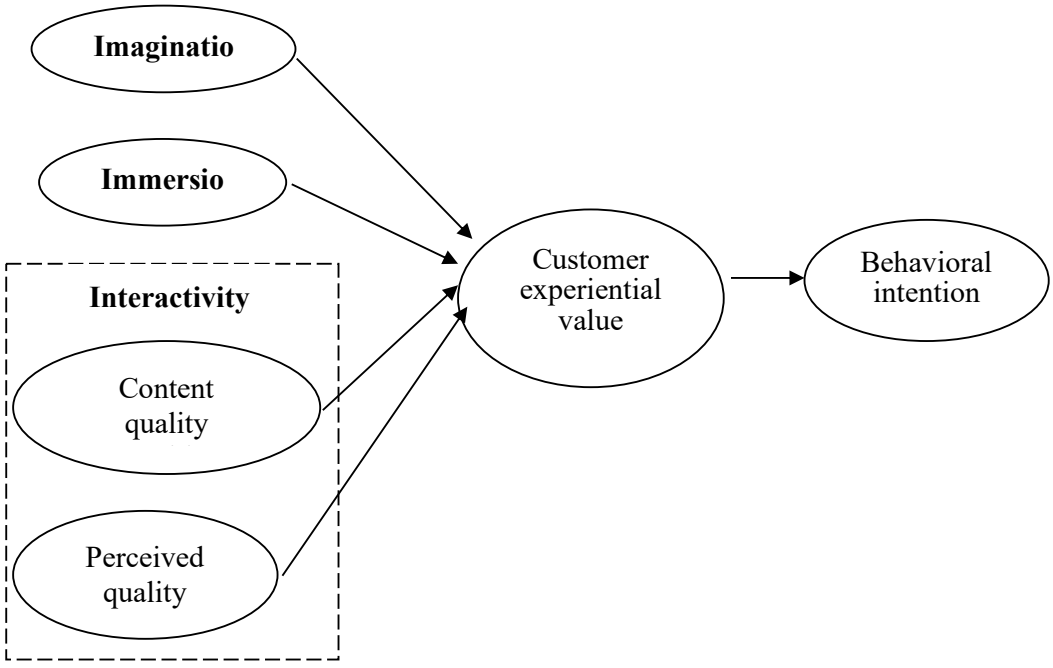


Figure 1 VR virtual tourism experience value model

Methods

Sample and data collection

This study adopts the parallel-translation method to ensure the applicability of the Chinese version of the questionnaire. A pretest was conducted based on the advice of one tourism scholar and three VR scholars to verify the conceptual, rather than the literal, equivalence between the measurement items in the original English and the items translated into Chinese. All the items are operationalized with a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). The subjects of this study are Gen Z, who were born between the late 1990s and early 2010s, respondents who have used VR virtual tours once or newly completed VR virtual travel. In this study, web-based questionnaires were distributed to collect data more conveniently and extensively. Gen Z completed the online questionnaire from social media platforms (e.g., Facebook and Instagram). In addition, we briefly described the research objectives and agreed to give each respondent a gift certificate (i.e., 100 New Taiwan dollars). These respondents answered questions related to each of the research constructs.

The questionnaires were distributed through the internet in 2021. A total of 559 samples were collected, of which 150 were invalid questionnaires. The number of valid questionnaires was 409. The characteristics of the sample are exhibited in Table 1. Among them, 148 males accounted for 36.2% of the total, while 261 females accounted for 63.8% of the total. Because this study was mainly composed of Gen Z, the age range was 20-25 years old. In the occupational part, 74.5% were students, and 232 people had the highest bachelor's degree, accounting for 56.7% of the total. In terms of income, most of the subjects were students. Finally, as many as 60.6% of the subjects were familiar with VR virtual tourism, and up to 80.9% of the subjects had used VR virtual tourism.

Harman's one-factor test was used to examine potential common method bias (CMB) (Podsakoff and Organ, 1986). Nine factors with eigenvalues greater than one accounted for 73.1% of the total variance, and the first factor accounted for 19.7% of the variance. Since a single factor did not emerge and one general factor did not account for most of the variance, CMB was unlikely to be a serious problem in this study.

Table 1. Demographic of the sample			
Item	**Category**	**N**	**Rate (%)**
Gender	Male	148	36.2%
	Female	261	63.8%
Age	20 and fewer	47	11.5%
	20 to 25	317	77.5%
	25 to 30	27	6.6%
	30 to 35	14	3.4%
	Over 35	4	1%
Occupation	Student	305	74.6%
	Office worker	104	25.4%
Education	Bachelor	232	56.7%
	Master	168	41.1%
	PhD	9	2.2%
Familiar with virtual tourism	Yes	248	60.6%
	No	161	39.4%
Experienced virtual tourism	Yes	331	80.9%
	No	78	19.1%

Data analysis and results

Measurement properties

Confirmatory factor analysis (CFA) was performed using AMOS 27.0. The results of the measurement model were Chi-square/degrees of freedom = 2.068, RMSEA = 0.05, GFI = 0.92, AGFI = 0.89, CFI = 0.97, SRMR = 0.03, and the overall measurement model had an acceptable fit. The internal consistency of the data was evaluated with three different measures: Cronbach's alpha, composite reliability (CR), and average extracted variance (AVE) (Fornell and Larcker, 1981). All the latent variables showed Cronbach's alpha values higher than the minimum threshold value of 0.70 suggested by Nunnally (1978), which indicates satisfactory internal consistency for confirmation purposes. To provide adequate internal consistency, the value of CR must be greater than 0.7 (Nunnally, 1978). The AVE values of all constructs were greater than 0.50, indicating that over 50% of the variance is explained by the measurement items. The results presented in Table 2 indicate internal consistencies, as both Cronbach's alpha and CR rates are above 0.7, and the AVE values were all above 0.5 for all constructs. Table 3 shows that the square root of the AVE extracted for each latent variable is all above the values for the correlations between the constructs. Hence, it can be assumed that the various constructs are reliable and have convergent validity.

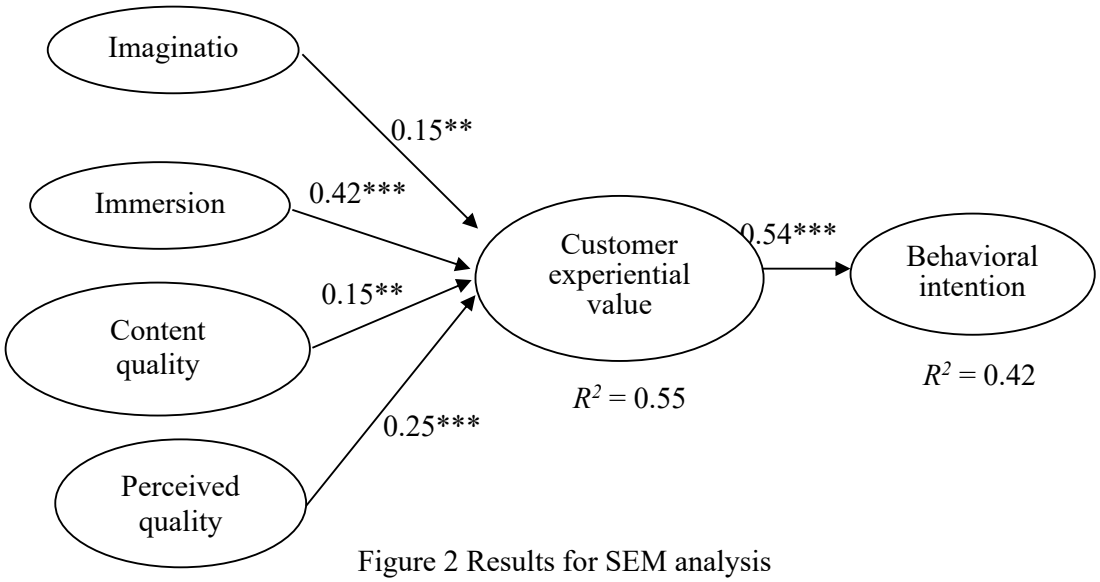


Figure 2 Results for SEM analysis

Hypothesis testing

The results showed that Chi-square/degrees of freedom = 2.23, RMSEA = 0.05, GFI = 0.91, AGFI = 0.89, CFI = 0.96, and SRMR = 0.03. This structural model had a good degree of fit. Figure 2 shows that the imagination characteristic of VR virtual tourism positively influences customer experiential value (β = 0.156, t = 2.713), supporting H1. The immersion characteristic of VR virtual tourism positively influences customer experiential value (β = 0.427, t = 5.959), supporting H2. The content quality cognition of the trip positively influences customer experiential value (β = 0.152, t = 3.010), supporting H3. The perceived quality of the trip positively influences customer experiential value (β = 0.253, t = 4.791), supporting H4. Customer experiential value positively influences behavioral intention (β = 0.54, t = 16.395), supporting H5. The predictive power of the research model is assessed by examining the explained variance for the endogenous constructs. Regarding the R^2 values, the characteristics of VR virtual tourism explain 55 percent of the variance in customer experiential value. In addition, customer experiential value explains 42 percent of the variance in behavioral intention.

Table 2. Measurement properties

Construct Name	Construct Identifier	Items	Factor Loading	Cronbach Alpha	Composite Reliability (CR)
Imagination	IMA	IMA1	0.89	0.89	0.88
		IMA2	0.86		
		IMA3	0.87		
		IMA4	0.88		
Immersion	IMM	IMM1	0.79	0.82	0.85
		IMM2	0.81		
		IMM3	0.84		
		IMM4	0.78		
Content quality cognition	CQC	CQC1	0.83	0.85	0.89
		CQC2	0.84		
		CQC3	0.86		
		CQC4	0.82		
Perceived quality	PQ	PQ1	0.86	0.88	0.88
		PQ2	0.83		
		PQ3	0.88		
		PQ4	0.85		
Consumer return on investment	CROI	CROI1	0.85	0.89	0.86
		CROI2	0.83		
		CROI3	0.85		
		CROI4	0.89		
Aesthetics	AE	AE1	0.87	0.88	0.87
		AE2	0.85		
		AE3	0.86		
		AE4	0.86		
Service excellence	SE	SE1	0.90	0.89	0.90
		SE2	0.89		
		SE3	0.88		
		SE4	0.88		
Playfulness	PF	PF1	0.92	0.93	0.91
		PF2	0.91		
		PF3	0.91		
		PF4	0.90		
Behavioral intention	BI	BI1	0.92	0.90	0.92
		BI2	0.94		
		BI3	0.89		
		BI4	0.91		
		BI5	0.90		

Table 3. Means, SD, correlations and average variance extracted (n = 409)												
Construct	Mean	SD	AVE	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMA (1)	4.14	0.66	0.55	**0.74**								
IMM (2)	3.99	0.72	0.63	.42**	**0.79**							
CQC (3)	3.72	0.71	0.54	.41**	.45**	**0.73**						
PQ (4)	3.89	0.67	0.79	.54**	.50**	.54**	**0.88**					
CROI (5)	3.79	0.64	0.70	.48**	.50**	.47**	.52**	**0.83**				
AE (6)	3.90	0.72	0.63	.53 **	.57 **	.60 **	.53 **	.43 **	**0.79**			
SE (7)	4.01	0.68	0.68	.55 **	.40 **	.60 **	.57 **	.29 **	.34 **	**0.82**		
PF (8)	4.22	0.56	0.70	.36 **	.36 **	.32 **	.44 **	.53 **	.41 **	.50 **	**0.83**	
BI (9)	4.15	0.61	0.77	.50 **	.36 **	.51 **	.54 **	.54 **	.44 **	.56 **	.32 **	**0.87**

Notes: a) Figures in shaded diagonal are values of the square root of the AVE; b) ** $p < 0.01$; c) IMA = imagination, IMM = immersion, CQC = content quality cognition, PQ = perceived quality, CROI = consumer return on investment, AE = aesthetics, SE = service excellence, PF = playfulness, BI = behavioral intention.

Discussion and conclusions

Theoretical contributions

This paper makes two important contributions. First, based on Burdea and Coiffet (2003), this study contributes to the literature by proposing a VR experiential value model including immersion, imagination, and both content quality cognition and perceived quality as VR virtual tourism characteristics behind customer experiential value. Second, whereas most prior research on customer experiential value has focused on analyzing consumer shopping behaviors, our study tests the theoretical model in the underinvestigated case of VR virtual tourism. Our research highlights that immersion, imagination and both content quality cognition and perceived quality help Gen Z attain good experiential value and that, in turn, customer experiential value has a positive impact on behavioral intention, thus confirming that customer experiential value contributes to the behavioral intention of VR virtual tourism and that the immersion characteristic of VR technology plays a central role in this experiential process.

Managerial implications

VR virtual tourism can be defined to enhance the tourism experience or another type of tourism,

which can benefit the tourism industry in terms of marketing, education, guided tours, and organizational benefits. This study explores the experience of Gen Z through smart devices such as VR wear when preparing for a destination tour: interactive experiences in virtual environments such as hotel rooms, attractions, museums, restaurants, etc., can make the prepurchase decision process easier. Search for information about the destination on the internet to obtain the opportunity to experience it in advance or decide whether the holiday or destination in question is suitable for them before purchasing. Therefore, Gen Z is interested in VR to provide an advanced experience of hotels, museums, and destinations before making travel decisions. For destination organizations and tourism enterprises, VR helps increase the attractiveness of destinations and tourism products (accommodation business, catering business, museums, historical sites, festivals, exhibitions, etc.) and reduces marketing and promotion costs. The advantages of these innovations ensure barrier-free tourism, for example, enabling tourists with disabilities and nondisabled tourists to experience tourism products in the same way as well as for promotion and marketing.

Limitations and future research

First, due to time constraints and the slow collection of samples, the questionnaire information was not provided as feedback after the first experience but was conducted by subjects who had experienced VR experience in the past and targeted subjects who had experienced travel VR experience (including registered members of the experience center). Recalling experience and filling out the questionnaires also indirectly resulted in too many invalid questionnaires. Second, expert interviews and user interviews can be added, which will help to enhance the authenticity and reference value of this study in terms of qualitative research.

Third, with the development of Metaverse and 5G technology, the content design of VR virtual tourism should be transformed into a guiding and public platform method. It is important to combine VR virtual tourism with the image of the on-site tourism destination and condense the value of the tour content to create the benefits of virtual and physical integration. This new

interactive experience will fundamentally change the way we interact, purchase and travel and attract the attention of consumers. Therefore, it is recommended that tourism operators provide an appropriate interaction mechanism through the online sharing interface to provide a social communication platform for virtual tourists to communicate with each other during travel, thereby reducing the loneliness of virtual tourists alone. In addition, the development of creative VR virtual tourism products (such as high-altitude jumping, diving, etc.) can expand the meaning and value of tourist attractions, which can be effectively transformed and applied in tourism to enhance customer participation .

References

Aaker, J. (1991). The negative attraction effect? A study of the attraction effect under judgment and choice. *ACR North American Advances.*

Bafadhal, A. S., and Hendrawan, M. R. (2019). Exploring the immersion and telepresence in gamified virtual tourism experience toward tourist's behavior. In *Annual International Conference of Business and Public Administration (AICoBPA 2018)* (pp. 53-56). Atlantis Press.

Beck, J., and Egger, R. (2018). Emotionalize me: Self-reporting and arousal measurements in virtual tourism environments. In Information and communication technologies in tourism 2018 (pp. 3-15). Springer, Cham.

Bhattacherjee, A., and Premkumar, G. (2004). Understanding changes in belief and attitude toward information technology usage: A theoretical model and longitudinal test. *MIS quarterly*, pp. 229-254.

Bilgihan, A., Barreda, A., Okumus, F., and Nusair, K. (2016). Consumer perception of knowledge-sharing in travel-related online social networks. *Tourism Management*, Vol. 52, 287-296.

Buehler, R., and McFarland, C. (2001). Intensity bias in affective forecasting: The role of temporal focus. *Personality and Social Psychology Bulletin*, Vol. 27 No. 11, pp. 1480-1493.

Buhalis, D., and Karatay, N. (2022). Mixed reality (MR) for Generation Z in cultural heritage tourism toward metaverse. *In ENTER22 e-Tourism Conference (pp. 16-27). Springer, Cham.*

Buhalis, D., Harwood, T., Bogicevic, V., Viglia, G., Beldona, S., and Hofacker, C. (2019). Technological disruptions in services: lessons from tourism and hospitality. *Journal of Service Management,* Vol. 30 No. 4, pp. 484-506.

Burdea, G. C., and Coiffet, P. (2003). Virtual reality technology. *John Wiley & Sons.*

Carter, T. J., and Gilovich, T. (2010). The relative relativity of material and experiential purchases. *Journal of Personality and Social Psychology*, Vol. 98 No. 1, pp. 146.

Chou, P. F., and Lu, C. S. (2009). Assessing service quality, switching costs and customer loyalty in home-delivery services in Taiwan. *Transport Reviews*, Vol. 29 No. 6, pp. 741-758.

Coyle, J. R., and Thorson, E. (2001). The effects of progressive levels of interactivity and vividness in web marketing sites. *Journal of Advertising,* Vol. 30 No. 3, pp. 65-77.

Del Bosque, I. A. R., San Martín, H., and Collado, J. (2006). The role of expectations in the consumer satisfaction formation process: Empirical evidence in the travel agency sector. *Tourism Management,* Vol. 27 No. 3, pp. 410-419.

Fornell, C. and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, Vol. 19, pp. 39–50.

Garvin, D. A. (1983). Can industry self-regulation work?. *California Management Review*, Vol. 25 No. 4, pp. 37-52.

Grace, D., and O'Cass, A. (2004). Examining service experiences and post-consumption evaluations. *Journal of Services Marketing,* Vol. 18 No. 6, pp. 450-461.

Holbrook, M. B., and Hirschman, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *Journal of Consumer Research*, Vol. 9 No. 2, pp. 132-140.

Holbrook, M. B., and Schindler, R. M. (1994). Age, sex, and attitude toward the past as predictors of consumers' aesthetic tastes for cultural products. *Journal of Marketing Research,* Vol. 31 No. 3, pp. 412-422.

Huang, M., Lu, Y., Sin, S. W., Seng-Pan, U., and Martins, R. P. (2016). A fully integrated digital LDO with coarse–fine-tuning and burst-mode operation. *IEEE Transactions on Circuits and Systems II: Express Briefs*, Vol. 63 No. 7, pp. 683-687.

Ioannides, D., and Gyimóthy, S. (2020). The COVID-19 crisis as an opportunity for escaping the unsustainable global tourism path. *Tourism Geographies,* Vol. 22 No. 3, pp. 624-632.

Itani, O. S., and Hollebeek, L. D. (2021). Light at the end of the tunnel: Visitors' virtual reality (versus in-person) attraction site tour-related behavioral intentions during and post-COVID-19. *Tourism Management*, Vol. 84, pp. 104290.

Kim, M. J., Lee, C. K., and Jung, T. (2020). Exploring consumer behavior in virtual reality tourism using an extended stimulus-organism-response model. *Journal of Travel Research,* Vol. 59 No, 1, pp. 69-89.

Klein, R. J., Nicholls, R. J., and Thomalla, F. (2003). Resilience to natural hazards: How useful is this concept?. *Global Environmental Change Part B: Environmental Hazards,* Vol. 5 No. 1, pp. 35-45.

Kwok, A. O., and Koh, S. G. (2021). COVID-19 and extended reality (XR). *Current Issues in Tourism,* Vol. 24 No. 14, pp. 1935-1940.

Kwon, J., and Lee, H. (2020). Why travel prolongs happiness: Longitudinal analysis using a latent growth model. *Tourism Management*, Vol. 76, pp. 103944.

Lee, S. A., Lee, M., and Jeong, M. (2021). The role of virtual reality on information sharing and seeking behaviors. *Journal of Hospitality and Tourism Management*, Vol. 46, pp. 215-223.

Lee, W., and Gretzel, U. (2012). Designing persuasive destination websites: A mental imagery processing perspective. *Tourism Management,* Vol. 33 No. 5, pp. 1270-1280.

Lemke, F., Clark, M., and Wilson, H. (2011). Customer experience quality: an exploration in business and consumer contexts using repertory grid technique. *Journal of the Academy of Marketing Science,* Vol. 39 No. 6, pp. 846-869.

Moon, J. W., and Kim, Y. G. (2001). Extending the TAM for a Worldwide-Web context. *Information & Management,* Vol. 38 No. 4, pp. 217-230.

Ntounis, N., Parker, C., Skinner, H., Steadman, C., and Warnaby, G. (2022). Tourism and hospitality industry resilience during the Covid-19 pandemic: Evidence from England. *Current Issues in Tourism*, Vol. 25 No. 1, pp. 46-59.

Nunnally, J. C. (1978), *Psychometric Theory (2nd ed.)*. New York: McGraw-Hill.

Overby, J. W., and Lee, E. J. (2006). The effects of utilitarian and hedonic online shopping value on consumer preference and intentions. *Journal of Business Research,* Vol. 59 No. 10-11, pp. 1160-1166.

Parasuraman, A., and Grewal, D. (2000). The impact of technology on the quality-value-loyalty chain: a research agenda. *Journal of the Academy of Marketing Science,* Vol. 28 No. 1, pp. 168-174.

Park, C. W., and MacInnis, D. J. (2006). What's in and what's out: Questions on the boundaries of the attitude construct. *Journal of Consumer Research*, Vol. 33 No. 1, pp. 16-18.

Persky, S., and Blascovich, J. (2007). Immersive virtual environments versus traditional platforms: Effects of violent and nonviolent video game play. *Media Psychology,* Vol. 10 No. 1, pp. 135-156.

Petrick, J. F. (2002). Development of a multidimensional scale for measuring the perceived value of a service. *Journal of Leisure Research,* Vol. 34 No. 2, pp. 119-134.

Podsakoff, P. M., and Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, Vol. 12 No. 4, pp. 531-544.

Pullman, M. E., and Gross, M. A. (2004). Ability of experience design elements to elicit emotions and loyalty behaviors. *Decision Sciences,* Vol. 35 No. 3, pp. 551-578.

Rahimizhian, S., Ozturen, A., and Ilkan, M. (2020). Emerging realm of 360-degree technology to promote tourism destination. *Technology in Society*, Vol. 63, pp. 101411.

Rajagopal, P., and Montgomery, N. V. (2011). I imagine, I experience, I like: The false experience effect. *Journal of Consumer Research,* Vol. 38 No. 3, pp. 578-594.

Schiopu, A. F., Hornoiu, R. I., Padurean, M. A., and Nica, A. M. (2021). Virus tinged? Exploring the facets of virtual reality use in tourism as a result of the COVID-19 pandemic. *Telematics and Informatics*, Vol. 60, pp. 101575.

Sherman, W. R., and Craig, A. B. (2018). Understanding virtual reality: Interface, application, and design. *Morgan Kaufmann.*

Steenkamp, J. B. E. (1990). Conceptual model of the quality perception process. *Journal of Business Research,* Vol. 21 No. 4, pp. 309-333.

Stepaniuk, K. (2015). The relation between destination image and social media user engagement–theoretical approach. *Procedia-Social and Behavioral Sciences*, Vol. 213, pp. 616-621.

Suh, K. S., and Chang, S. (2006). User interfaces and consumer perceptions of online stores: The role of telepresence. *Behavior & Information Technology,* Vol. 25 No. 2, pp. 99-113.

Sweeney, J. C., and Soutar, G. N. (2001). Consumer perceived value: The development of a multiple item scale. *Journal of Retailing,* Vol. 77 No. 2, pp. 203-220.

Tussyadiah, I. P., Wang, D., Jung, T. H., and Tom Dieck, M. C. (2018). Virtual reality, presence, and attitude change: Empirical evidence from tourism. *Tourism Management,* Vol. 66, pp. 140-154.

Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., and Schlesinger, L. A. (2009). Customer experience creation: Determinants, dynamics and management strategies. *Journal of Retailing,* Vol. 85 No. 1, pp. 31-41.

Wang, Y. J., Minor, M. S., and Wei, J. (2011). Aesthetics and the online shopping environment: Understanding consumer responses. *Journal of Retailing*, Vol. 87 No. 1, pp. 46-58.

Wei, W., Torres, E., and Hua, N. (2016). Improving consumer commitment through the integration of self-service technologies: a transcendent consumer experience perspective. *International Journal of Hospitality Management*, Vol. 59, pp. 105-115.

Yoo, J., and Kim, M. (2014). The effects of online product presentation on consumer responses: A mental imagery perspective. *Journal of Business Research*, Vol. 67 No. 11, pp. 2464-2472.

Yung, R., and Khoo-Lattimore, C. (2019). New realities: a systematic literature review on virtual reality and augmented reality in tourism research. *Current Issues in Tourism,* Vol. 22 No. 17, pp. 2056-2081.

Zeithaml, V. A., Berry, L. L., and Parasuraman, A. (1996). The behavioral consequences of service quality. *Journal of Marketing*, Vol. 60 No. 2, pp. 31-46.

Abstracts

Accounting, Finance, Economics

A Leveraged Bootstrap Causality approach to investigate relationships among 11 GICS stock market sectors and the Covid-19 cases after the declaration of Pandemic

Abstract

Dr. Howard Lee [1]

1. *Bloomsburg University of PA*

This research investigates the causal relationships between the performance of the 11 GICS stock market sectors and the case number that are infected by the Covid-19 after the WHO Declares COVID-19 a Pandemic. The relationships were investigated using the Granger causality test with a leveraged bootstrap approach. The results indicate that at the beginning of the event occurred, the returns of 11 GICS stock market sectors were influenced by case number of Covid-19 highly. After 6 months, the case number of Covid-19 lost some influence, but only on the Healthcare sector. Our findings indicate that the declaration of the Covid-19 Pandemic has great influence on the entire stock market at the first 6months. As time goes by, the event was lost the attention by the investors and the stock market was no longer driven by it. This research provides perspective for building investment portfolios after a global event occurs.

An Examination of Operating Foreign Exchange Exposure

Abstract

Dr. Augustine Arize [1], Dr. John Barkoulas [2], Dr. John Malindretos [3], Dr. Alex Panayides [4], Dr. Chuanqian Zhang [4]

1. *Texas A&M University – Commerce,* ***2.*** *Georgia southern,* ***3.*** *william Paterson University,* ***4.*** *William Paterson University*

Operating foreign exchange exposure is the change of the currency value affecting the value of the subsidiary in another country, if let us say, the parent is in the US. The change in currency value has to be unexpected, so that the risk managers will not have taken steps to protect the firm from loss. The concept of operating foreign exchange exposure is one of long term. Consequently, contractual hedges generally will be appropriate to protect the firm from probable harm only in the short run, defined as a year or two. More appropriate hedging techniques are diversification of operations, sales and financing, matching currency cash flows, currency swaps and risk sharing. In this study, we report the actual methods finance practitioners use to hedge.

Are Commercial Lending Decisions Affected by Severity of Internal Control Material Weaknesses and Tolerance for Ambiguity?

Abstract

Dr. Arnold Schneider [1]

1. *Georgia Institute of Technology*

This study examines whether the severity of a material internal control weakness disclosure impacts lending decisions. In addition, this study also investigates whether loan officers' tolerance for ambiguity affects lending decisions. An experiment was conducted with loan officers, who provided risk assessments and lending probabilities for a hypothetical borrower. One independent variable, a material weakness in the borrower's internal controls, was manipulated using two levels – more severe versus less severe. The other independent variable, tolerance for ambiguity, was measured with a commonly used scale for this personality characteristic. Findings indicate that more severe material internal control weaknesses lead to higher risk assessments by lenders and lower probabilities of granting lines of credit. Also, lenders having a low tolerance for ambiguity provide higher risk assessments and lower probabilities of granting lines of credit than do lenders having a high tolerance for ambiguity.

Corporate Social Responsibility Reporting Improvements and Mandated CO2 Reductions in the Northeast United States: New Evidence on Positive Outcomes from an Unexpected Symbiotic Relationship

Abstract

Dr. A J Stagliano [1]

1. *Saint Joseph's University*

Voluntary financial disclosure theorists argue that it is in the firm's best interest to report information about costly sustainability initiatives. For those who hold to this reporting theory, the quantity/quality of disclosure should be positively related to performance. Legitimacy theorists propose, differently, that a firm presents itself in a way that society wants it to be. Those proposing that legitimacy theory is the more appropriate view posit that there is will be an inverse relationship between climate change performance and financial reporting. Current empirical evidence might assist us in assessing which of these two competing theories more correctly is reflected in contemporary accounting and reporting.

This research focuses on whether government mandates to enhance sustainability generate an additional positive externality by stimulating greater financial disclosure about environmental performance. We study fossil-fuel fired electric generating plants in the northeastern U.S. to determine whether when faced with a requirement to curtail CO_2 emissions impacted facility owners' financial reports show better disclosure regarding environmental performance.

The study years are 2008 through 2015. Emissions information comes from the U.S. Environmental Protection Agency. Financial disclosure intensity regarding environmental matters is assessed through a content analysis of annual filings made to the U. S. Securities and Exchange Commission.

With a matched-sample design, we find significant improvement in environmental performance subsequent to implementation of a CO_2 cap and strong indications that financial disclosure regarding corporate social responsibility for this sustainability initiative was enhanced. Mandating sustainability initiatives appears to stimulate greater social responsibility reporting to stockholders.

Educating business schools students in financial metrics of climate impact in turbulent business environments

Abstract

Dr. Carmen Quirvan [1]

1. *The University of Rhode Island*

Educating business schools students in financial metrics of climate impact
in turbulent business environments.
CARMEN QUIRVAN, Ph.D.
THE UNIVERSITY OF RHODE ISLAND
ABSTRACT
There is a necessity to include the environmental studies into the education of business schools' students to study the impact of businesses into the biodiversity. We need to educate these students to perform research in environmental accounting for finding the correct financial metrics regarding the visible impact of the climate change.
Business schools should improve their curriculum educating students for the future not only in biodiversity but also for the financial system in turbulent business environments.

Examining “Boilerplate” CAM Reporting with NLP

Abstract

Dr. Nathan Slavin [1]***, Dr. Jianing Fang*** [2]***, Mr. David Fisher*** [3]

1. *Hofstra University,* ***2.*** *Kean University,* ***3.*** *USC*

The Public Company Accounting Oversight Board (PCAOB) has significantly expanded the auditor’s report in an unqualified opinion for large accelerated filers for fiscal years ending June 30, 2019. The objective of the regulation was to provide the investor with additional information regarding areas where the auditor encountered subjective, complex and high risk areas in completing the audit examination. The objective of these CAM disclosures was to alter the pass/fail evaluations of the client’s annual report where special attention would be communicated on the most difficult and contentious areas. Both the PCAOB and Jay Clayton, the former chairman of the Securities and Exchange Commission warned against using a “boilerplate” approach in conforming to these new CAM disclosures where auditors would merely report the same CAMs between reporting years without addressing any modifications that could have affected the disclosures.

Comparing the practice of “boilerplate” reporting for CAMs is extremely time consuming in determining whether accountants are practicing “boilerplate” reporting. We have examined an efficient tool in the programming area of Natural Language Processing (NLP) to determine the existence of “boilerplate” practice using members of the DOW 30 companies for the fiscal years of 2019 and 2020 as a potential and efficient measurement of a “boilerplate” practice. The measurements used in the NLP procedures were compared to the individual evaluations of the CAM reports in determining its veracity and validity as a “boilerplate” context detector.

How to Measure the Financial Impact of ISO 14001 Certification: Introducing New Financial Metrics

Abstract

Dr. Tom Geurts [1]***, Dr. Beate Klingenberg*** [2]

1. *Bucknell University,* ***2.*** *FOM Hochschule für Ökonomie und Management*

In previous research the authors established that widely employed financial ratios should not be used for measuring the impact of ISO 14001 certification on the financial performance of firms. This conclusion was drawn after reviewing the literature on this topic, which has produced inconsistent results. It was found that these inconsistent results were the consequence of differences in the definition and measurement of financial performance, the methodology for firm matching, and the ultimate research design and data analysis. Using a new dataset and after removing the ambiguity surrounding these four issues, it was concluded that when employing the aforementioned traditional financial ratios combined with the most reliable methodology available, there is no significant impact of ISO 14001 certification on financial performance. However, there are three ways to interpret this result:

- There is indeed no impact of certification on financial performance,
- The methodology that is being used, is not the best methodology, and
- ROA and ROE are not the best financial ratios to measure financial performance.

Since the event study has traditionally been a reliable methodology for this type of research (Barber and Lyon, 1996; Corbett et al., 2005) and the answer of no impact is intellectually unsatisfactory, the authors will next explore finding better financial ratios, namely ones that are not subject to the influence of cash flows that are unaffected by the implementation of ISO 14001. Exploring the framework of cost deployment (Giovando et al., 2020), we will present the findings of this research in this paper.

Impact of Rising Diesel Prices and Truck Driver Availability on Food Transportation and Distribution

Abstract

Ms. Fuqin Zhou [1], ***Dr. Aichih (Jasmine) Chang*** [1], ***Dr. Jim (Junmin) Shi*** [1]

1. *New Jersey Institute of Technology*

Since the outbreak of the COVID-19 pandemic, the food transportation and distribution has been critically impacted by the rising oil price and trucker shortage, and the food price has been soaring with the increasing transportation cost. To investigate the relationship between transportation costs of food and food prices, this paper proposes a systematical regression to unveil the impact of the surging fuel prices and truck driver availability. We collect a panel data consisting of diesel price, truck driver wage, and food price (e.g., bacon, banana, beef, bread, chicken, coffee, cookie, egg, flour, milk, orange, rice, tomato, etc.) since the outbreak of COVID-19 (i.e., January 2020), and adopts the fixed effects regression to specifically examine the increase of food price stemmed from the surging fuel price and truck driver wage, while controlling the factors from food demand (e.g., GDP). It has been widely proved that oil supply disruptions may cause economic turbulences from the three oil supply shocks in the past two decades. This study contributes to the literature by focusing on the food industry to examine how the food price has been impacted jointly by the rising fuel price and truck driver shortage in the past two years. Additionally, we further investigate the issues about the long-lasting trucker shortages, current and prospective economic trends and uncertainties, and the fuel price trend, and their impacts on the food price.

Measure Annual Report Narratives Disclosure: Empirical Evidence from Disclosure of COVID-19 Information in Jordan

Abstract

Dr. Suzan Abed [1], ***Prof. Tih Koon Tan*** [1], ***Dr. Batao Chen*** [1]

1. *University of the District of Columbia*

This paper aims to investigate empirically the common alternative methods of measuring annual report narratives. Two alternative methods are employed, an un-weighted disclosure index and automated content analysis using word counts. The authors investigate the extent of voluntary disclosures of COVID-19 information of 60 Industrial Jordanian companies for the year 2021. Descriptive analysis and correlation matrix are employed to measure the correlation between the employed methods to measure the extent of voluntary disclosure of COVID-19 information. The results reveal that while the alternative methods of voluntary disclosure of COVID-19 information are highly correlated, important significant differences do nevertheless emerge. It appears important to measure volume rather than simply the existence or non-existence of each type of disclosure. Overall, the optimal method is the automated content analysis by word count, which enables authors to code several annual reports at the same time. This paper contributes to the extant literature on voluntary disclosure of COVID-19 information by reporting important differences among alternative disclosure methods. However, the decision regarding whether this should be a computerized or a manual content analysis appears not to be driven by differences in the resulting measures. Rather, the choice is the outcome of a trade-off between the time involved in setting up coding rules for computerized analysis versus the time saved undertaking the analysis itself

Product Advertising and Financial Analyst Forecasts

Abstract

Dr. Seung Won Lee [1]

1. *Penn State Harrisburg*

We examine whether financial analysts discover new information from firms' product advertising activities. Our analysis uncovers that analysts provide more informative and accurate sales forecasts as a firm's product advertising intensity increases. A difference-in-differences test and an instrumental variable analysis identify the causal relationship. Additionally, the effect of advertising on analyst sales forecast quality is stronger when a firm has a greater coverage from industry-expert analysts, higher sales volatilities, and register new trademarks. Overall, our evidence suggests that analysts are able to discover new information from advertising activities to improve sales forecast quality.

RESILIENT COMPANIES AFTER FINANCIAL CRISIS, PANDEMIC AND ENVIRONMENTAL REPORTING

Abstract

Dr. Carmen Quirvan [1]

1. *The University of Rhode Island*

RESILIENT COMPANIES AFTER FINANCIAL CRISIS, PANDEMIC AND
ENVIRONMENTAL REPORTING
CARMEN QUIRVAN, Ph.D.
THE UNIVERSITY OF RHODE ISLAND
ABSTRACT
We have had several main economic events in this century, the 2008 financial crisis in the United States (Quirvan, et al., 2014), after that the pandemic came as a type of health crisis; recently the awareness of environmental issues related to air, water, biodiversity in US and around the world.
We can see examples of resilient companies such as the utilities industry in the US especially on the north, in which these companies face high levels of capital spending due of their own requirements of updating old structures and the transition to a clean economy. The recent news about the carbon border tax that the European Union is implementing and the very recent meeting in Canada for finding consensus regarding climate damages among rich and poor nations; all of these are environmental issues to be aware. This is a time in which managers of companies need to solve many challenges facing rising costs in operating such as wages, materials and supplies, and debt. Another key factor is inflation which affect companies. In addition, public companies need to follow the requirements established by the SEC of the US to inform about their climate impact because of their operation; but companies' metrics are correct?
Our world is facing many challenges in which the only solution is the cooperation among nations, companies, and peoples to have strong companies that can go beyond of these tests.

Supplies' Upstreamness and Investment Inefficiency: The Effect of Auditors from Downstream

Abstract

Prof. SUNGSOO KIM [1]

1. *Rutgers University - Camden*

ABSTRACT

We examine whether suppliers' investment inefficiencies increase with their upstreamness in the supply chain, and whether the sharing of an auditor with their major customer can mitigate such effect. Using a comprehensive dataset of multi-tier supply chains, we report that firms further upstream in the supply chain exhibit more investment inefficiency. However, this inefficiency can be mitigated when these upstream suppliers engage the same auditor as their major customer. Furthermore, we find that this mitigating effect is stronger when this auditor is a specialist in the major customer's industry or has longer tenure with the major customer. Our findings suggest that such mitigating effect operates at least in part through an information sharing channel. Taken together, we provide a novel consequence of the upstreamness in the supply chain, namely, investment inefficiency, and demonstrates that auditors with downstream expertise can serve as information intermediaries to mitigate this consequence

The New Economy and its Impact on the Flux of Digitalized Information in the United StatesAmerica

Abstract

Prof. Billy Bataille [1]

1. *Seton Hall University*

Abstract

To study the dynamic process of transitioning to a new digitalized American economy, 3 variables: education, employment and internet access are used to create a transition index to the new economy. Statistical data reveals a complex bidirectional relationship linking the aforementioned variables to describe the impacts of technological shocks initiating adaptive behavior from actors involved.

Introduction

I construct an index of transition from the industrial age to the new economy to measure multidimensional changes that create a technological transition affecting the American population and exacerbating the digital divide and inequality, ans its effect on education and the labor market, focusing on California, Florida, Indiana and South Dakota.

Relevant Issues

I) The above-mentioned sectors are intertwined and their associated variables correlated.

II) Actors interact using adaptive strategies maximizing their utility.

III) Because some of the actors' decisions are not directly observable, proxy variables (i.e.: GDP, employment, education, wage) are used.

Methodology

I) Define the transition to the new economy using the relative GDP of different economic sectors.

II) Identify the effects of the transition.

III) Collect statistical data on U.S. states to quantify the relevant proxy variables, taking into account the correlation existing between them.

IV) Establish a linear dynamic model to model the interaction between the actors dynamically adapting their optimizing strategy to the changing environment. I directly/indirectly observe state variables representing the actors' behavior and measuring outcomes.

V) Statistical data helps to calibrate the model and develop a transition index to analyze the transition to the new economy.

Timing Value versus Growth

Abstract

Dr. Pawan Madhogarhia [1]

1. *York College of Pennsylvania*

Growth has outperformed value over the last several years. Is value dead? Value tends to outperform growth over longer horizons. This paper explores if investors can identify whether growth or value portfolios are expected to outperform at a certain time. Value and long term earnings growth spreads are used to address this research question. Zero investment portfolios yielded positive returns for different growth and value portfolios using returns with different frequencies. At the time of writing this paper, it appears that value portfolios are predicted to outperform growth portfolios in the next decade.

Big Data, Analytics, and Knowledge Management

Business Intelligence Platform Development for Monitoring Bus Operation Risk Indicators

Abstract

Ms. Xiaotong (Claire) Ding [1], ***Prof. John Maleyeff*** [2], ***Mr. Frans Valk*** [1]

1. *Massachusetts Bay Transportation Authority,* ***2.*** *Boston University*

With buses that travel over 20 million miles per year, the MBTA is concerning about passenger safety around the greater Boston region. A project was initiated to develop a business intelligence (BI) platform to monitor bus incidents, such as collision with vehicles, fire and smoke, onboard customer injury, etc. The project started by analyzing various information sources to identify influential factors using logistic, polynomial, and lasso regression to rank features having a high impact on bus accidents. Centroid-based methods, distribution model-based clustering, and hierarchical algorithms were used to categorize bus operators' driving actions including number of extreme events, excessive speed, Max G force, etc. An interactive BI platform was designed after validating statistical significance. These features included bus operators' driving behavior, incident records, customer complaints about incidents, bus vehicle engineering information, and real-time bus performance tracking data. The platform was implemented using SplunkCloud. Development challenges included combining various data sources with different time spans, examining significant relationships for feature reduction, creating visualizations for real-time monitoring of bus operator and route performance, and making the system user-friendly.

Digital Marketing Conversion on Social Media

Abstract

Dr. Rae Yule Kim [1]

1. *Montclair State University*

The asset value of advertising is often difficult to measure because the effect of advertising is intangible. This paper utilizes conditional inference tree (CT) methods to develop prediction models for advertising conversion to estimate the value of advertising on business performance. The ultimate goal for digital marketers is to increase customer conversion, such as making a purchase or subscribing to a service. We utilize customer conversion records of digital ads on Facebook to examine factors that best predict advertising conversion. Spending more does not always lead to more conversion events. Conversion tends to increase exponentially if marketers make more than a certain amount of investment in the ad campaign. The number of clicks and investment amount together best predict advertising conversion.

Exploring Expectancy Theory for Business Analysts: The Option of Failure

Abstract

Dr. Ellen Belitzky [1]

1. *University of New Haven*

Practical teaching experience in industry and academia has rendered a reasonable sample to infer a common business analytics mindset: Failure is unacceptable. Expectancy theory in Vroom's "Work and motivation" (1964) explains analyst motivation in terms of performance rewards. In volatile business environments, there is risk when business analysts are compensated to understand and present data informing enterprise solutions, yet this work mindset may become part of an organization problem. Expectancy theory applied in business analytics must also reward analysts who fail thereby motivating improved organizational decision-making. Industry might decouple analyst incentives from business growth alone and could consider loss prevention and risk mitigation in reward structure. Business analytics educators have a responsibility to teach students who believe they must always succeed about the cost of failing to fail. Failure can be defined as the inability to find complete and reliable data, solve a problem, or pass an examination. Failure should not result in academic or career tragedy. The opposite is true as failing some challenges is exactly what will happen when working in industry. During an educational program, students can learn from this academic experience and work together with classmates and instructors to develop technical skills, professional habits, and personal coping strategies. Academia is great practice for building resilience. Given many business analytics instructors have been successful in careers requiring their own analytics skills and may not be trained educators, the best lesson to teach analysts may be that failure could be a viable option.

Introduction to Business Analytics as a required core course in the undergraduate business program at the University of Hartford: Discussion on coverage of topics

Abstract

Dr. Bharat Kolluri [1], ***Dr. Ning Jia*** [1]

1. *University of Hartford*

At present, we have two required core quantitative courses in our undergraduate business program. Both of these together cover the traditional topics in business statistics from descriptive statistics to regression analysis with excel applications through all the topics. In addition, some topics on optimization are also covered. During the last five-year period, Barney school of business conducted a survey of our alumni and board of advisors, to collect information on what more we could do to prepare our students for the workforce. Programming/software skills came on at the top, such as knowledge of Python, R and some of SQL, Tableaux, power BI, etc in addition to Excel. As Business Analytics becomes the most popular major, and business analytics skills become more commonly required for all business majors entering the job market, an introduction to business analytics course will very possibly become a core course. This comes at a price of combining the existing two beginning quantitative course and the third current elective, an introductory business analytics course into two required core quantitative courses for all majors. This purpose of this presentation is to discuss and seek suggestions and input on topics of coverage, especially in the Introduction to business analytics course.

Knowledge Sharing Incentives in AI-Enabled Knowledge Management

Abstract

Dr. Shankar Sundaresan [1], ***Dr. Zuopeng Zhang*** [2]

1. *Rutgers University - Camden,* ***2.*** *University of North Florida*

We develop an analytical model of AI-enabled knowledge management in organizations. Analyzing this model, we explore the design of a reward scheme to motivate knowledge workers' input in an AI system to improve its quality and the design of a payment structure to induce workers' application of the system to produce an enhanced output. Our research lays the foundation for future research in AI-enabled knowledge management and provides insights for practitioners to improve the efficacy of AI implementations.

Mapping of Multi-Year Decision Making Strategies for Climate Change Uncertainties

Abstract

Dr. Bryan Higgs [1], ***Ms. Karla Dimitri*** [1]

1. *University of the District of Columbia*

The uncertainties associated with climate change predictions create a strain on decision makers of coastal communities where they must plan for a wide variety of potential future scenarios of sea level rise ranging from 0.28 meters to 1.01 meters by the year 2100. Decisions made in the present not only affect immediate outcomes but influence the availability of choices in the future, thus decision makers must make long-term plans that best protect coastal communities in the present and in the future. This research effort explores the development of a multi-year decision making framework based in game theory where future decisions can be based on future framework states and potential future outcomes. This framework represents how decision makers can choose to shift or even abandon strategies based on the current level of knowledge with the level of knowledge changing over the years as new data is collected. One major aspect of the framework is that decision makers can learn when critical decisions must be made thus balancing the balancing the prospects of the clarification given by new knowledge against the risks of the delay of critical decisions The key outcomes of the game theory framework are the identification of robust multi-year decision strategies that best position coastal communities to be able to adapt to potential future climate scenarios while they are evolving. The major benefit of this research is the ability to avoid the situation where adaptation becomes impossible which will result in catastrophic losses for a coastal community.

Meme Stocks: Social media meet gamified markets

Abstract

Ms. Isha Thombre [1], ***Ms. Mahika Nair*** [1], ***Prof. Amit Das*** [2]

1. *School of Interwoven Arts & Sciences, Krea University,* ***2.*** *IFMR Graduate School of Business, Krea University*

In January 2021, a group of mostly amateur investors came together on an Internet forum called r/wallstreetbets to coordinate their purchase of the stock of videogame retailer GameStop, forcing its price sharply upwards, to the detriment of those who were short selling the stock. The squeeze to cover the short positions sent the stock price even further up, followed by a crash. Using data from the user posts on r/wallstreetbets forum for over a month, and the price data of GameStop stock over the same time-frame, we examine the relation between the sentiment embodied in the forum postings and contemporaneous and delayed movements in the stock price. Sentiment analysis of each day's corpus of forum posts is conducted using state-of-the-art natural language processing (NLP) libraries. The relationship of the series of daily sentiment measures to the series of stock prices and trading volumes (with varying lags) is assessed using correlation, controlling for such factors as the movement of the overall market. While the effect of internet activity, such as tweets and blogs, on stock prices has been studied for over a decade now, the GameStop case provides a clearer opportunity to observe this relationship and the emergence of the class of "meme stocks" in today's turbulent financial markets. Meme stocks represent an endogenous source of volatility arising from the interaction of widespread social media access and new-age discount brokerages that gamify stock trading.

Overhearing Hushed Voices: Using Unobtrusive Methods to Uncover Work Sentiments of People with Epilepsy

Abstract

Dr. Asha Rao [1], ***Dr. Surendra Sarnikar*** [1]

1. *California State University East Bay*

One in four Americans has a disability. Yet, in diversity discussions, disability takes a backstage to issues of race and gender despite the underemployment and discrimination faced by people with disabilities. We focus on one disability, to understand the issues faced by people with epilepsy (PWE) in the workplace. Both the illness and associated stigma create distress, reducing self-efficacy and the quality of life for PWE. We use unobtrusive data mining of anonymous discussions by PWE to hear concerns that are difficult to disclose at work. By understanding their work-related motivations, concerns, and emotions, and any perceptions of discrimination, organizations can better support people with disabilities.

Data came from a public community forum on the Epilepsy Foundation from posts by adults and caregivers on managing and living with epilepsy. We collected 7142 discussions posted over 16 years from 2004 through 2020. We used Linguistic Inquiry and Word Count (LIWC), a computerized text analysis tool that uses a dictionary-based approach for analyzing linguistic and psychological constructs expressed in text data. Topic modeling and computerized text analysis is being increasingly used in management research and has the potential to advance management theory by helping detect novel and emergent phenomena based on an analysis of text data. We found that their discussions on work were more positive compared to non-work posts. They were analytical, emotional and achievement oriented. However, compared to benchmark corpora, PWE were more anxious, health centered and tentative. We suggest managers examine the psychological components of accommodation to support PWE.

Sequential Pattern Analysis Under Brand Loyalty Context

Abstract

Dr. Jin Fang [1], ***Dr. Hanxi Sun*** [2], ***Dr. Junhee Kim*** [3]

1. *Clark University,* ***2.*** *Purdue University,* ***3.*** *California State University Stanislaus*

This paper sheds new light on sequential purchasing behavior and brand prioritization in a loyalty program partnership, providing novel insights into loyalty program literature. Our network-based model captures significant temporal structures and provides meaningful clustering solutions to group consumers. Moreover, our model uses a Hyperlink-Induced Topic Search algorithm (HITS) to prioritize brands within a loyalty program partnership. According to this study, sequential purchasing orders differ among consumer groups.

Using GIS and network modeling for hay export

Abstract

Dr. EunSu Lee [1]

1. *New Jersey City University*

This study investigates sourcing strategies for livestock feedstock and transportation plans from the United States to Korea. The Korean hay market has been steadily growing in transactions for the last decade due to increased beef consumption. Korean feedstock importer has been operating a local factory in the United States to improve the procurement process and acquire sustainable supply. The study estimates the level of hay inventory and pasture production to promote hay, grassland and agricultural exports. This study utilizes the Geographic Information System (GIS) to predict production in the fields and to estimate effective intermodal freight transport routes. The study uses mixed integer programming (MIP) to find the best transportation options and suppliers based on the existing feedstock factory. The study utilizes geographic information systems (GIS) to predict production farms and estimate effective intermodal freight shipping routes. Preliminary results of the study show that the location of ports and rail terminals are one of the critical factors of production site. Preliminary results from this study show that states close to the port benefit and can utilize empty containers.

Using learning analytics to improve students' performance

Abstract

Dr. Youqin Pan [1], ***Prof. Jian Gu*** [1]

1. *Salem state university*

Learning analytics has attracted more attentions since more analytics and statistics courses are offered online due to the effect of covid-19. Learning data are collected by different courseware platforms. The purpose of this study is to examine how faculty make use of the Hawkes' learning data in a statistics course to monitor and predict student performance. Moreover, several issues and concerns regarding the use of learning analytics in higher education were discussed.

Using Predictive Analytics to Forecast Player Performance in Poker Tournaments

Abstract

Dr. Michael Paz [1], ***Dr. Mikhail Sher*** [2], ***Dr. Robert Scott III*** [2]

1. *Cornell University,* ***2.*** *Monmouth University*

The use of performance analytics has drastically changed the landscape of North American sports in recent decades. In this paper, we utilize predictive analytics and hand collected empirical data to build models that forecast player performance in high-stakes poker tournaments. We consider early predictors of success within multi-day poker tournaments and conduct probabilistic analyses of performance milestones and related outcomes based on early-stage tournament performance. These analyses have important implications both from a game theory standpoint as well for practical applications of our findings.

Cyber Security, IT, and Emerging Technologies

A SEEDS' Secure Federated Framework for Global Food Security and Climate Intelligence Data Repositories

Abstract

Dr. Benjamin Branch [1], ***Mr. Thomas Glenn*** [2], ***Mr. Micheal Sullivan*** [2], ***Mr. Jarvis Green*** [3], ***Mr. Bill Wright*** [4], ***Dr. Neset Hikmet*** [5], ***Mr. Luigi Zuccarelli*** [6], ***Dr. Charles Zelek*** [7]

1. *QSBG, SEEDS Institute,* ***2.*** *SEEDS Institute, Global Supply Chain Security SME,* ***3.*** *Green Reef Foundation, Oceans 97, Commercial Fisheries and Food Distribution SME,* ***4.*** *Red Hat, Head of Edge and AI; Enterprise Neurosystem Lead,* ***5.*** *University of South Carolina, Professor of Integrated Information Technology,* ***6.*** *Red Hat, Principal Software Engineer,* ***7.*** *CEO, Climate Smart Business Solutions, LLC, Global decarbonization and energy SME*

This Federated Framework for Global Food Security and Climate Intelligence concept is to better Global food supply chains, and climate change intelligence challenges which are further exacerbated by siloed systems, lack of data integrity, and concerns over information privacy. Food and climate intelligence require expedited, robust security, and a federated data framework may aid decision-making, data integrity, and sustainable data repositories. Training on issues of zero trust, GitOps, and open-source software deployment protocols (e.g. Kubernetes) are needed for system sustainability. As the Enterprise Neurosystem (EN) community seeks to normalize data-sharing of global climate and bio-indicators for research, correlation with this community may aid in the development of secure data workflows and repositories. Ongoing EN work suggests that full and joint lifecycle training and prototyping through coordinated hackathons and workshops could foster interoperability, standardization, and innovative solution development.

Specifically, this effort will conceptualize training in an open-source, single-node installation of OKD. The Community Distribution of Kubernetes that powers Red Hat OpenShift at okd.io offers a fundamental cybersecurity discussion for edge support concepts.

SEEDS Institute is a 501c3 committed to strategic competition, DEIA, technology, education and workforce, and sustainability solutions in global strategic industries and geographies. SEEDS proposes the following priorities in Kubernetes prototyping of such use cases:

Intentional, open-source pipeline development with emerging technologies in a full lifecycle development, prototyping effort.

More effective and customized data rendering that empowers a unified response (outreach and in-reach, with predictive models and milestones) is needed to ensure data reliability, provenance, and reproducibility thresholds.

Cultural Values and Artificial Intelligence

Abstract

Dr. Subhasish Dasgupta [1], ***Dr. Babita Gupta*** [2], ***Dr. Yuan Xue*** [3]

1. *George Washington Universtiy,* ***2.*** *California State University Monterey Bay,* ***3.*** *Elizabeth City State University*

Individuals and organizations are developing new applications that use artificial intelligence. Little research has been done to examine factors that influence use of artificial intelligence-based systems. In this research, we explore cultural values that shape individual attitudes toward intelligent applications, which in turn, impacts its use and adoption. We first define what is artificial intelligence and examine its components. Then we review recent advances in artificial intelligence and its applications. After examining advances in artificial intelligence, we focus on some important questions related to individuals' acceptance and use of artificial intelligence (AI). How do people feel about using/adopting AI applications? Do individuals trust AI? Are they willing to provide information to an AI application? Are individuals willing to let AI do some jobs that they do now? From the cultural perspective, we examine how an individual's cultural values influence their attitude towards AI. Their attitude could impact their use of AI-based applications.

CYBERCHONDRY (CSS-12) SEVERITY SCALE: A SPECIFIC APPLICATION TO HEALTH INFORMATION

Abstract

Dr. Neset Hikmet [1]***, Prof. Gamze YORGANCIOĞLU TARCAN*** [2]***, Prof. Meltem SAYGILI*** [3]***, Prof. Ayca Karahan*** [4]***, Mr. Nihat Baris Sebik*** [5]

1. *University of South Carolina,* ***2.*** *Hacettepe University,* ***3.*** *Kırıkkale University,* ***4.*** *Pamukkale University,* ***5.*** *Republic of Türkiye Ministry of Health*

Using the Cyberchondria Severity Scale (CSS-12) developed by McElroy et al. (2019) and administered to IT workers in the field of health informatics, the goal of this study is to evaluate whether cyberchondria differs based on socio-demographic characteristics. Participating in the study were one hundred (n=113) computer engineers, computer programmers, computer technicians, and software engineers from Turkish public hospitals. There is a wide range of factor loads for the items on the scale (0.536 to 0.768), and these factors explain 72.37 percent of the overall variation. According to the confirmatory component analysis, the CSS-12 scale consists of 4 factors and 12 items (Chi-square/freedom value = 2.064; RMSEA = 0.071; NFI = 0.916; CFI = 0.954; GFI = 0.929; AGFI = 0.915). It was determined that the reliability of the scale was 0.80 based on its internal consistency. Using appropriate statistical parametric or non-parametric tests based on the assumption of normality of the research data and homogeneity of variances, the objective of the study was anticipated to be reached.

Examining the Relationship Between Social Media and Cryptocurrencies

Abstract

Dr. Subhasish Dasgupta [1], ***Dr. Yuan Xue*** [2], ***Dr. Michael Chuang*** [3]

1. *George Washington Universtiy,* ***2.*** *Elizabeth City State University,* ***3.*** *University of Illinois*

Cryptocurrencies have received a lot of attention in recent years. Bitcoin perhaps one of the most popular cryptocurrencies has changed the landscape of financial and currency markets. As we know, financial markets are greatly influenced by customer cognition, sentiment, and behavior. Cryptocurrencies are no different. Customers and traders of cryptocurrencies have frequented social media platforms to exchange views, ideas, and expertise in the field. Our study focuses on the effect of social cognition, sentiment, and behavior on cryptocurrency trading, and vice versa. We analyze user-generated data from social media sites that discuss cryptocurrency trading by using data mining techniques such as sentiment analysis to examine this relationship.

High-level vulnerabilities: The role of age in information security awareness and attitudes

Abstract

Dr. Gregory Lyon [1]

1. *Georgetown University*

Despite the immense growth of internet and digital device usage among older individuals—who make up a growing proportion of the workforce and who often hold high-level decision-making positions in firms—empirical research on the cyber behavior of older individuals is limited. This study draws on two large national surveys of adults to examine the relationship between age and objective cyber knowledge and attitudes. Using a two representative U.S. surveys and weighted logistic regression models, the study estimates whether and to what extent age affects whether individuals can accurately identify a set of core secure cyber practices and the extent to which age affects attitudes toward federal information security policy. The results indicate that older individuals lack knowledge of secure cyber practices and support augmenting federal cyber infrastructure. The largest knowledge gaps between older and younger individuals are found for some of the most important secure cyber practices including multifactor authentication, one of the most effective mechanisms to defend against cyberattacks. The findings have important implications for our understanding of information systems, older individuals' cyber behavior, and user vulnerability in an increasingly cyber-connected society.

Industry 4.0 Technologies in Maritime Industry: Systematic Literature Review and Future Research Direction for ESG Investing

Abstract

Dr. Leo Hong [1], ***Dr. Se Kyu Kim*** [2], ***Ms. Kara Liu*** [3], ***Dr. Douglas Hales*** [3]

1. *Millersville University,* ***2.*** *Department of Sustainability Management-Inha University,* ***3.*** *University of Rhode Island*

Industry 4.0 technologies play an important role in achieving many sustainable development goals such as preserving biodiversity in the maritime industry. Industry 4.0 technologies have the potential to enhance human efforts to protect the environment and conserve resources by detecting energy emission reductions, CO2 removal, and monitoring extreme weather conditions. Industry 4.0 technologies also allow investors to collect and analyze more information than ever when accounting for environmental, social, and governance risks and growth opportunities (Pournader et al., 2021). These capabilities are helpful for ESG investing, which reflects the growing sensitivity of investors and how companies should operate in their decision-making. Moreover, investment managers are experiencing pressure to measure ESG criteria in their portfolios and the lack of data makes it hard to assess the long-term risks of maritime-related problems.

However, research to date has been primarily restricted to engineering or technical disciplines, focusing on problem-solving issues. Theory development regarding the use of Industry 4.0 technologies for a sustainable maritime industry is still in the nascent stage. To elaborate on our focus, we evaluated the linkage of Industry 4.0 technologies characteristics using the 'ESG' framework model. Given the novelty of the area, we hope to contribute to the literature and practice by providing the necessary foundation of how different technologies benefit the maritime industry in terms of convincing skeptical investors that the organization's actions are sincere. We review the literature related to using technologies' applications in the maritime industry and provide future directions.

The Effect of Perceived Blockchain Technology Benefits on ESG Disclosure Trustworthiness in South Korea

Abstract

Dr. Se Kyu Kim [1], Dr. Leo Hong [2], Dr. Jong Dae Kim [3]

1. *Department of Sustainability Management-Inha University,* ***2.*** *College of Business - Millersville University,* ***3.*** *College of Business Administration - Inha University*

This study is an empirical study of the perceived benefits of blockchain technology on ESG disclosure trustworthiness. Survey data collected from 361 professional investors in South Korea are used to validate our arguments. The study employs comprehensive and rigorous procedures to test the hypotheses developed. Leveraging the Unified Theory of Acceptance and Use of Technology view of blockchain, we establish a theoretical framework for hypothesis testing. We use factor analysis followed by hierarchical regression analysis.

The findings present that blockchain's benefits (1) efficiency pursuit, (2) trust, and (3) risk avoidance have positive impacts on the reliability of ESG disclosure through the mediation variable: the usefulness of blockchain technology. In addition, the level of ESG knowledge positively moderates the relationship between blockchain's efficiency, trust, and ESG disclosure trustworthiness. It also positively moderates the usefulness of blockchain technology and ESG disclosure trustworthiness.

In conclusion, we present that the perceived benefits of blockchain technology positively impact the reliability of ESG disclosure among professional investors in South Korea. This suggests that increasing investors' positive perception of blockchain benefits can resolve ethical conflicts regarding the credibility of ESG disclosure which directly links to corporate ESG performance. The findings are applicable to the service industry sector in the future. The dynamics and implementation of blockchain technology in ESG are worthwhile topics to pursue in future research.

The Escape Room: Powering Up Your Spreadsheet Using Advanced Design, Formulas, and Functions as Applied to the Accounting Discipline

Abstract

Dr. Mark Law [1], ***Dr. Cassandra Bennett*** [1]

1. *Bloomsburg University of PA*

This paper describes a project-based educational resource to help students develop a better understanding of the interrelated accounting practice and theory as well as introduce and reinforce advanced Microsoft Excel functions, formulas, and spreadsheet design as applied to an “escape room”. This paper provides examples of spreadsheet layouts, advanced formulas and functions, and accounting fundamentals and has practical implications for educators teaching accounting classes incorporating information technology applications.

Decision Making: Public Administration and Policy

A Demographic Characterization of Supplemental Nutritional Assistance Program (SNAP) Recipients with Machine Learning Models

Abstract

Dr. Mehmet Kilinc [1]***, Mrs. Montserrat Avila Acosta*** [2]

1. *Le Moyne College,* ***2.*** *The University at Albany, SUNY*

The purpose of this study is to identify the factors that influence families' likelihood of receiving Supplemental Nutritional Assistance Program (SNAP) benefits. SNAP is the largest federal program providing nutritional assistance to low-income families in the United States. We use data from the Survey of Income and Program Participation (SIPP). We will compare the performance of a logistic regression model with machine learning algorithms, including random forest, neural networks, and gradient boosting machines, to determine which demographic and other family characteristics affect their likelihood of receiving SNAP benefits. Specifically, we examine whether family composition, size, race, ethnicity, education, presence of children, presence of elderly, presence of members with disabilities, and the safety of their neighborhood, make families more or less likely to receive SNAP benefits.

A Mixed Approach for Highway-Rail Garde Crossings Risk Analysis Considering Crash Likelihood and Severity

Abstract

Dr. Amin Keramati [1], ***Dr. Pan Lu*** [2], ***Prof. Afrooz Moatari-Kazerouni*** [1]

1. *Assistant Professor, School of Business Adminstration, Widener University,* ***2.*** *Associate Professor, North Dakota State University*

The highway-rail grade crossing (HRGC) is a specific spatial location where rail track and road intersect at grade level. The HRGC crashes usually have relatively severe results. In addition, traffic delays on both the railway and the roadway can considerably extend the economic consequences of crashes at HRGCs. Identifying systematic methods is essential to ensure that federal and state funds for highway-rail grade crossing improvement projects are allocated to locations and crossings at higher risk of an accident. Hazard index and collision prediction formula techniques are the most general prioritization approaches for ranking highway-rail grade crossings. While the hazard index is used to estimate a value that ranks crossings in relative terms (the higher the quantified index, the more hazardous the crossing), the collision prediction formula (prediction model) is utilized to quantify the predicted crash frequency or severity. A few research projects and state DOTs used a hybrid model considering both a crash frequency and a hazard index approach. In this study, two hazard-ranking models are proposed; one is based on the crash likelihood estimated by using competing risk model, and the other is a "hybrid accident prediction model-hazard index" based on collision severity likelihoods using both comporting risk model and Analytic Hierarchy Process (AHP). Finally, the risk analysis is conducted using the risk matrix technique to integrate the results of both hazard-ranking approaches and classify grade crossings considering both crash frequency and severity likelihood simultaneously.

Air Connectivity Policy and Socioeconomic Factors: Evidence from Sub Saharan Africa

Abstract

Prof. Soheil Sibdari [1]

1. *University of Massachusetts, Dartmouth*

We use an international aviation data set to examine whether higher aircraft capacity in the trans-Atlantic markets is associated with higher trade and tourism. We use international and domestic data sets between years 2009 and 2019 and focus on a selected courtiers in Sub-Saharan Africa and Southern Europe. Our results help policy makers, commercial airlines, and international organizations to make better capacity investment and route planning.

Application of Qualitative Judgments Based on Ordinal Model in Public Policy Decision Making

Abstract

Prof. Alexander Mechitov [1], ***Prof. Helen Moshkovich*** [1]

1. *University of Montevallo*

Empirical studies indicate that in many complex decision making problems the decision makers (political leaders, administrators, managers, experts) primarily formulate public policy tasks in qualitative form, which makes natural to use ordinal scales. The proposed approach divides the decision process in two steps: first, experts evaluate different alternatives using ordinal scales, and second, policy makers define their policy by using pairwise comparisons to define Joint Ordinal Scale. The application of Ordinal Model makes it possible to facilitate the description of public policy tasks, to structure collective decision making, and finally to determine the best decision among available options. The approach is supported by an example which illustrate application of this approach for public policy task.

COVID-19 Vaccine Line Jumpers and Social Welfare

Abstract

Dr. Jimmy Chen [1]

1. *Bucknell University*

During the initial rollout, the COVID-19 vaccine was only available to certain priority groups, and anyone not in a priority group had to wait their turn to be vaccinated.

Reportedly, many requesters from non-priority groups cut in the line to receive the vaccine before it was their turn. These requesters are often referred to as “vaccine line jumpers.”

Some argue that vaccines should be strictly reserved for priority groups.

Others argue that jumpers can make good use of any spare vaccine vials before they expire and that our society as a whole will benefit from more people being vaccinated, no matter what group they belong to.

In light of this debate, this study investigates whether vaccination sites should ban jumpers altogether or allow some jumpers to the vaccination site when balancing wait times for requesters and vials.

In particular, we propose a line jumper admission heuristic based on the number of jumpers allowed to minimize the time requesters and vaccines spend in the queuing system.

We use a queue system to simulate the operation of the vaccine site and the random arrivals of requesters and vaccines.

Digitization and Digital Transformation in Recent Years: Are South African Rural Areas Improving?

Abstract

Mrs. Sylvia Azwidohwi Siphugu [1]

1. *University of Johannesburg*

Sylvia Siphugu[1]

[1] Johannesburg Business School, University of Johannesburg

Abstract

Digital transformation of urban areas has been a focus specialty around the world, lately. Neglect, limited access to digitalization and other basic amenities are the major criteria for defining rural communities. The world value digital transformation has a projected spending of $1.8 trillion and $2.8 trillion globally in 2022 and 2025, respectively. Recent studies conducted for the digitization of basic services mainly focused on urban areas with fewer literature for the rural communities led by tribal chiefs. This trend seems to be prevalent in South Africa wherein preliminary literature review suggests that studies done in digital transformation are mostly in urban areas. Hence, this study aims to assess digitization improvement and digital transformation in recent years whether South African rural areas improving or not. Data from existing literature are qualitatively analysed. Mixed method was used to evaluate the data outcomes. Results shows that digital transformation has been a major focus in urban areas whereas rural areas remain inert and non-progressive. This could be due to lack of digital skills and poor advanced information technologies in rural areas led by tribal chiefs. The result from this study showed that the utilization of AI in basic services administration could accelerate and ease the bottleneck hassles faced by the tribal chiefs in carrying out their basic administration services as leaders. The study aims to identify gaps for future research.

Keywords: Digital Transformation and Digitization, Tribal Chiefs, Rural Areas, Basic Services

The Role of Online Influencers and Data Quality in the Customer Decision-making Process

Abstract

Dr. Saeed Tabar [1]

1. *Assistant professor, Miller College of Business, Ball State University*

In today's electronic business environment, online influencers play an important role in customers' decision-making process. They collect and analyze information and provide the processed information to customers to facilitate their decision-making. When customers purchase an item, they go through a five-stage process including need recognition, information search, evaluation of alternatives, purchase decision, and post-purchase behavior to make the final decision. Online influencers can impact this process especially by collecting information on various products and brands, evaluating different alternatives, ranking items in the choice set, and finally picking the best item that can fulfill customers' needs. The quality of the information provided by online influencers is important because inaccurate information could result in inaccurate decisions and financial loss. Data quality has four dimensions including intrinsic, contextual, representational, and accessibility with their respective subdimensions.

Therefore, the objectives of this project are threefold: 1-study the role of online influencers in new product adoption and diffusion, 2-investigate the customer purchase decision-making process and how online influencers can facilitate this process, 3-explore the dimensions and subdimensions of data quality and see which dimensions can contribute to customer product adoptions. To fulfill the mentioned objectives, a comprehensive literature review on main research constructs such as opinion leadership, customer decision-making process, and data quality will be conducted. Following the literature review, hypotheses on the interrelationship of these constructs will be developed. The research instrument will be adopted from respective past projects. Finally, data will be collected and then tested to check the proposed hypotheses.

DSS, Machine Learning, and Artificial Intelligence

A Comparative Study of EES Technologies using DEA Models

Abstract

Dr. Ahmet Akgun [1]***, Dr. Mehmet Yildirim*** [2]

1. *Pennsylvania Western University,* ***2.*** *Wichita State University*

To avoid the environmental damages of fossil fuels from electricity generation, policymakers aim to increase the share of renewable energy in power generation. In recent years, renewable energy investments have increased significantly to minimize the negative impact of fossil fuels. However, the efficient usage of renewable energy is tough to accomplish due to the intermittent capacity of renewable energy. Electrical energy storage (EES) units are one way to increase the efficient utilization of renewable energy to address intermittency concerns. As a result, the selection of the EES technology is very critical due to the requirements of the different applications. This study proposes a data envelopment analysis (DEA) model to select the best technologies among other EES units. Input and output-oriented DEA models have been implemented under four different returns to-scale assumptions: Constant Returns to Scale (CRS), Variable Returns to Scale (VRS), Increasing Returns to Scale (IRS) and Decreasing Returns to Scale (DRS). 10 different EES technologies have been selected to compare the relative efficiencies and 63 EES units considered in the model. The results show that the proposed model provides effective results and this method can be used to compare the EES technologies for different applications with better data.

A DEA-Based DSS for Food Shoppers

Abstract

Dr. Christine Pitocco [1], ***Dr. Thomas Sexton*** [1]

1. *Stony Brook University*

We use Data Envelopment Analysis to build a multicriteria decision support system for food shoppers. The model measures the performance of each food item along each criterion (nutrient) relative to other food items in the same food group, identifies food items that lie on the frontier, and measures the distance that each food item lies from the frontier with respect to each criterion. These factor performances are critical for those who shop for people with nutritional constraints (sugar, salt, calories, etc.). The model uses data from the U.S. Department of Agriculture.

A Two-Stage Machine Learning Method to Predict Cross-Buying Decisions

Abstract

Dr. Mehmet Kilinc [1]***, Mr. Robert Rohrhirsch*** [2]

1. *Le Moyne College,* ***2.*** *MEAG*

In this paper, a new approach to predicting cross-buying behavior in customers is proposed. This involves using machine learning, feature engineering, and interpretation techniques to create a two-stage prediction framework. The first stage involves training a complex, black-box classification model using cross-validation and tuning of parameters. In the second stage, the top ten most important predictors from the first model are used to create a simple, interpretable rule-based model. The effectiveness of this approach is demonstrated using a publicly available dataset, and the results indicate that the rule-based model performs just as well as the complex model.

CorMiDEA: A Novel Feature Selection Algorithm for Improved Learning with High Dimension and Low Sample Size

Abstract

Mr. Mojtaba Talaei-Khoei [1], ***Dr. Asil Oztekin*** [1], ***Prof. Yao Chen*** [1]

1. *University of Massachusetts Lowell*

Machine learning algorithms tend to perform poorly when the number of features is high relative to the sample size, a phenomenon known as the *curse of dimensionality*. One of the ways to tackle poor learning in high-dimension, low-sample space is to reduce the number of features to lessen the dimensionality of the data, hence decreasing the noise and possibly increasing the machine learning algorithm's performance. Correlation coefficient and mutual information are used in numerous algorithms as the mechanisms for feature selection. Feature selection based on mutual information is one of the best dimensionality reduction techniques to create precise and robust machine learning models. On the other hand, data envelopment analysis (DEA) is a powerful non-parametric modeling technique to assess the efficiency of decision-making units (DMUs). Input-oriented Constant Returns to Scale (CRS) multiplier was hybridized along with mutual information and correlation coefficient to create three integrated feature selection algorithms: 1) CorDEA, with correlation as the DEA output; 2) MiDEA, with mutual information as the DEA output; and 3) CorMiDEA, with correlation and mutual information as two outputs of the DEA. The efficiency of the features was measured and used to rank the features. A randomly selected low sample of the benchmark machine learning dataset from the UCI archive was used. Incorporating DEA as a feature ranking mechanism significantly increased the models' performance compared to mutual information and correlation coefficient alone. The methodology and case study validates the novel integrated feature selection algorithm, i.e., CorMiDEA, as it outperforms the other studied models.

How Blockchain PaaS and 3PLs can help disrupted logistics network

Abstract

Ms. Lanqing Du [1]***, Dr. Jinwook Lee*** [1]

1. *Drexel University*

Labor constraints have had a significant impact on traditional transportation and supply chain management operations since COVID. Reports and statistics from the government and major consulting companies imply such influence would be projected to impair the transportation and logistics industry in the long run. When faced with interruption and its associated ripple impact, third-party logistics (3PL) services are critical logistical fulfillment options when supported by cost-efficient technology enablers. In this research, evaluating the importance of information sharing, we investigate related decision-making problems in a disrupted logistics network, where nodes represent facilities with varying degrees of disruption and edges indicate a shipment volume between two nodes. Key findings from numerical experiments on real-world data will be presented.

Machine Learning in Supply Chain Management

Abstract

Mrs. Toyin Asonibare [1], ***Dr. Abirami Radhakrishnan*** [2]

1. *Morgan State,* ***2.*** *Morgan State University*

Digitalization has led to the generation of Big Data in the field of supply chain management (SCM). Although Big Data offers numerous benefits, its potential value cannot be realized unless it is evaluated for trends and insights that supply chain (SC) managers may utilize to make timely interventions. Big Data Analytics is the process of obtaining useful insights from large amounts of data using advanced analytic methods. Among the approaches, machine learning has gained significance. Several algorithms can be utilized to glean insights from SC related Big Data with minimal human interaction. However, SC managers and data scientists face the difficulty of utilizing machine learning algorithms to generate business value. Consequently, we are conducting a thorough literature review on the various machine learning techniques and their applicability to SCM. The findings of this study shows that machine learning techniques such as Neural Networks, Artificial Neural Networks, Support Vector Machine, Logistics regression, Decision Tree, Extreme Learning Machine, Naive Bayes Classifier, k-Nearest Neighbor, Random Forest, Ensemble Algorithms, and K-means are applied in different areas of SCM such as: risk assessment, supplier selection, fraud detection, demand forecasting, predicting backorders of products and reorder point assessment.

SENTIMENT, ATTITUDE, AND INTENTION TOWARD USING ARTIFICIAL INTELLIGENCE TECHNOLOGY

Abstract

Dr. Sung Shim [1]

1. *Seton Hall University*

This study examines the sentiment, attitude, and intention toward using artificial intelligence (AI) technology, using data collected from a survey of 220 business professionals. Regarding the sentiment toward using AI technology, this study uses the Positive and Negative Affect Schedule (PANAS) developed by researchers from the University of Minnesota and Southern Methodist University. The PANAS consists of two 10-item scales to measure both positive and negative affect. The ten items that describe positive affect are interested, excited, strong, enthusiastic, proud, alert, inspired, determined, attentive, and active. The ten negative items that describe negative affect are: distressed, upset, guilty, scared, hostile, irritable, ashamed, nervous, jittery, and afraid. Regarding the attitude and intention toward using AI, this study uses the constructs and items used by many studies on the technology acceptance model. The data will be analyzed in three stages: (1) descriptive analysis of the study constructs and items of positive affect, negative affect, attitude, and intention toward using AI technology, (2) factor analysis of the study constructs of positive affect, negative affect, attitude, and intention toward using AI technology, and (3) structural equation modeling of the study constructs of positive affect, negative affect, sentiment, attitude, and intention toward using AI technology to measure and analyze the relationships of the constructs. The results of this study could provide timely and valuable insights into the perceptions of business professionals toward using AI technology.

THE ROLE OF ARTIFICIAL INTELLIGENCE IN ADULT LUNG CANCER DETECTION

Abstract

Mrs. Sukaynah Al Haji [1]***, Dr. Roderick Lee*** [1]

1. *Penn State Harrisburg*

Abstract

Lung cancer remains the leading cause of cancer deaths worldwide. Symptoms usually appear after the cancer has spread, making the disease difficult to treat. Current medical practices present several challenges as there is a shortage of physicians in the United States. Also, the main screening methods produce a high-false positivity rate. In this research-in-progress, current literature is examined with the goal to determine the effects of implementing Artificial Intelligence (AI) to pathological screenings and physician decision making. The literature is thematically analyzed and summarized to conclude the role AI plays in lung nodule detection. AI is determined to greatly improve physicians' lung cancer detection and diagnosis ability. However, gaps in research and practice persist. This study concludes with implications for research and practice.

Keywords: artificial intelligence (AI), lung cancer, lung cancer detection, adults, radiographic scans.

Towards Interpretable Machine Learning

Abstract

Mr. Soumyabrata Basu [1], ***Prof. Amit Das*** [1]

1. *IFMR Graduate School of Business, Krea University*

Even when machine learning systems display a high level of performance, the inability of domain experts to understand how they work – how they arrive at their answers – can inhibit the adoption of such systems in organizations. This opacity leads to a lingering suspicion that the performance on one set of data set might not be repeated on the next set of data. Opacity also makes it difficult to justify the system to various stakeholders, and raises issues of accountability when the system fails, i.e. it makes incorrect decisions.

We apply three different machine learning algorithms – logistic regression, decision trees, and neural networks – to the problem of predicting corporate bankruptcy with a large set of independent variables. All three algorithms perform comparably in terms of classification performance (F1-score, Matthews correlation coefficient, and area under the ROC), though they weight the independent variables differently to arrive at their conclusions.

The three models, each with performance data and the set of top predictors (using Shapley values), are presented to domain experts: university professors of finance. We ask them to assess each model in terms of the theoretical justification for its selected predictors. A model that uses more justifiable predictors is more interpretable than others. We compare the interpretability ordering of the algorithms to their performance order. Finally, we run the three algorithms using only variables shortlisted by the human experts; these models represent a trade-off between performance and interpretability.

Education, Curriculum, and Cases

A Data Envelopment Analysis Approach to Rank Administrations of United States Academic Institutions

Abstract

Dr. Richard Muszynski III [1]

1. *Wilkes University*

Campus closures and mergers are becoming more frequent since the beginning of 2012. These events have caused university administrations to propose strategic plans that focus on recruiting prospective students, and many university administrations allocated time and resources to addressing the issue of a consistency downward trajectory in student enrollment. Prospective students also desire to know how each academic institution performs. We solve a data envelopment analysis (DEA) problem that informs students and university personnel about the performance and efficiency of each academic institution by identifying top administrations within higher education. Other university administrations may use this ranking to help better position themselves as a successful competitive academic institution. To identify the best university administrations, we select several variables that we believe help determine which academic institutions are performing well versus ones that show major concerns. These variables include net present value for fifteen years, net present value for thirty years, average monthly search volume divided by full time enrollment, and cost-to-value. We perform empirical analysis that involves comparisons between different categorizations of academic institutions including the type of academic institution, the size of each institution, and the state that each university resides.

A Knowledge Management Solution for Business Analytics Education

Abstract

Dr. Ellen Belitzky [1], ***Ms. Leeparani Parlapalli*** [1]

1. *University of New Haven*

With online teaching prominent throughout the Covid-19 pandemic, learning management systems (LMS) became repositories for course knowledge. For business analytics students from diverse backgrounds, leveraging LMS provided hands-on ease of practice with tools to solve problems they would encounter in their future careers. Rethinking course requirements involved risk; students might not correlate concepts and tools usage due to lack of insights and experience in real-world business operations. To mitigate risk, a mandate for certificate of completion LMS upload from a prerequisite online tutorial prior to the course helped students establish a common baseline of knowledge so that during the course they could focus on honing their skills and maximizing the value of the learning activities. Using the LMS as a discussion forum to supplement classroom work helped leverage student diversity, encourage the less verbal students to share, and showcase a wide range of viewpoints than would have been possible via LMS or on ground class each independently. Given limited class time, using the LMS enabled every student to participate. Having LMS discussion as a course requirement ensured every opinion was heard. Students generally had a positive view on assigned class work through the LMS with automated grading given immediate feedback on formative assessments. The educational impact of this course design was to create skilled business analysts completing degree programs who will work in industry and be prepared to build business resilience and enable successful commercial results.

Keywords: knowledge management, business analytics, analytics education, business resilience, post-pandemic education

AN EXAMINATION OF THE ROLE OF INCLUSIVE, EFFECTIVE, AND INNOVATIVE PEDAGOGICAL PRACTICES IN STEM EDUCATION

Abstract

Mr. Chandra Niredi [1]***, Dr. Roderick Lee*** [1]

1. *Penn State Harrisburg*

Throughout history, race and socioeconomic status have remained two of the most reliable predictors of STEM student success. Despite millions invested in programs and services to support student success, the opportunity gaps in the retention, persistence, and graduation rates for structurally disadvantaged students such as first-generation, low-income, students of color, and numerically underrepresented students in STEM persists (Boucher et al., 2021; Student Experience Project, 2022). Research shows that students learn more and are more likely to succeed when faculty leverage research-proven pedagogical practices such as active learning and culturally responsive teaching. However, the systemic adoption of inclusive pedagogical practices remains elusive, leaving unrealized their full benefits for equity and excellence in STEM education. Therefore, the purpose of this research-in-progress is to examine the efficacy of two inclusive pedagogical practices in STEM education: active learning and culturally responsive teaching. We conclude with implications for research and practice.

Keywords: Active learning, culturally responsive teaching, STEM education.

Analysis of the factors impacting enrollment in higher education institutions using Interpretive Structural Modeling

Abstract

Dr. Arsalan Paleshi [1], ***Dr. Aman Gupta*** [2]

1. *York College of Pennsylvania,* ***2.*** *Embry-Riddle Aeronautical University Worldwide*

The higher education institutions are forecasted to witness a major drop in enrollment starting the year 2025. The impact of this phenomenon is so severe that it has been referred to as enrollment cliff. Regional universities and colleges will be impacted the most by this reduction. The enrollment rates will continue to be low even in 2030s. The regional higher education institutions need to develop appropriate strategies to prepare for this major challenge. In particular it is important to know the factors that impact the high school students' decision in choosing a college for their studies. Furthermore, what are the interrelationship of these factors and which ones are more important. In this research, we use Interpretive Structural Modeling (ISM) methodology to understand the relationships of these factors and provide recommendations to these institutions to improve their enrollments during the hard years to come.

Analytics Skill Building: Leveraging R in the First Statistics Course

Abstract

Dr. Linda Boardman Liu [1]

1. *Boston College*

Over the last several years, we have revised and updated the undergraduate business school statistics course in myriad ways with a focus on creating connections between the statistics curriculum to the rest of the business school curriculum, to professional experiences, and to everyday applications of statistical analysis and analytical thinking. A particular focus has been on creating an environment that ensures students are business analytics ready, with modern computational skills, computer literacy, and some basic coding skills.

We introduced R/RStudio to the introductory statistics course with a variety of supporting course design elements including programming labs, tutorials, curated data sets, and other teaching innovations. These structural and pedagogical changes have resulted in significant benefits to student learning, engagement, and skill acquisition.
Analysis of student learning shows that regardless of prior education or training in statistics or programming, students report significant gains in analytical skills. For knowledge acquisition, students derive more learning from experiential and tactile activities that require critical thinking and problem-solving.
This early engagement with applied, real-world analytics methodologies has resulted in an increased level of interest and enrollment in the various business analytics courses that build upon the statistics core course, including management science, business intelligence, marketing analytics, and risk & simulation, as well as a heightened engagement in analytics-focused professional endeavors post-graduation.

Career Advisement Using Career Portfolio Planning Program: The Impact on Academic Performance, Retention Rate, and Job Placement Outcome

Abstract

Dr. Kihwan Kim [1]

1. *Kean University*

The prospect of finding a nice job after graduation is one of the strongest motivators that make students retain their program. As a tool that guides students' career development, we devised a "Career Portfolio Planning Program" which lists the requirements of knowledge, skill, and abilities for obtaining a specific job. Under this program, if the students complete the listed requirements, they are highly likely to get the job they wish for. According to the expectancy theory and goal-setting theory, when people have specific goals and believe they can achieve those goals, they are more likely to be motivated. In the context of college, if students have clear career goals and a strong belief in achieving the goals, they are more likely to retain their education. The program will serve as a tool that provides the expectancy and specific goals for achieving a better career. We will recruit 200 freshman students and divide them into the control and manipulation groups. The manipulation group receives advisement from Career Portfolio Planning Program while the control group receives regular advisement. The study proposes that this program would positively impact academic performance, student retention, and job placement rates.

Classroom Innovation: Teaching Undergraduate Business Analytics Using Publicly Available, Free, Real World Data

Abstract

Prof. Allison Miller [1]

1. *Georgia College & State Univerity*

Over the course of two academic years (8 classes, 228 students), the format for an undergraduate Business Analytics Course was changed. This new approach went from a more traditional use of textbook data bases and simple format solved problems to one which required them to "hunt and kill" their own databases by using free, publicly available data bases and performing analyses utilizing various business analytics tools. The idea was that in today's "Big Data" business world, students would benefit from having course projects they interacted with data which may not be as "clean" or easy to interpret, and this would give them experience with data from a more realistic standpoint.

This paper summarizes the results of this format change both from the standpoint of the students and the instructor. For both students and the instructor the challenges and benefits of the new versus traditional approach are considered and discussed including the implications of adopting this new approach for teaching the course in the future.

Creating a Culture of Connectedness and Collaboration in Asynchronous Learning

Abstract

Dr. Kellyann Kowalski [1], ***Dr. Jennifer Swanson*** [2]

1. *University of Massachusetts, Dartmouth,* ***2.*** *Stonehill College*

During Covid-19 most college courses were moved online, but were often taught synchronously. Being synchronous it was easier to reproduce the active learning classroom environment where students collaborate with each other and are able to develop connections throughout the semester. Now that most colleges are back to teaching courses on campus, most online courses are being taught in an asynchronous manner. But how do we teach asynchronously and still create a culture of connectedness and collaboration?

In our presentation we will address this general question and other more specific questions such as: What are the challenges faced by instructors and students in terms of developing connections? How can courses be designed to engage all students in a meaningful way with each other and the instructor? What types of tools and techniques are most effective to get and keep students connected throughout the entire course? How do you balance students' need for flexibility (which is often the reason they are taking asynchronous online courses) and the need for connectedness and collaboration?

We will discuss how creating a culture of connectedness and collaboration is something that needs to be done throughout the course, with activities planned for the beginning, middle, and end of the course. We need to go beyond the introduce yourself discussion board to create more meaningful interactions and connections. After a brief presentation, we will leave time to open up the discussion to the audience to share their experiences and to brainstorm other possible solutions.

References available upon request.

Data Analysis of the Teaching of Data Analytics Courses

Abstract

Dr. Derald Wentzien [1]

1. *Delaware State University*

Graduates with excellent data analysis are in high demand by many companies associated with many different types of businesses. It is very important to ensure that business students are educated and trained with these skills so that they can be effective employees. The purpose of this research was to use the data analysis process to identify the desired skills and select the methodologies that will effectively teach the skills. The six-step data analysis process promoted by Google was used to perform the analysis.

The six steps of the data analysis process used in this research was ask, prepare, process, analyze, share, and act. A word count of the desired data analysis skills requested in job postings indicated that the ability to process data and communicate the results to shareholders were the most requested skills. Suggestions from business faculty members and recent graduates indicated that students struggle with the ability to perform basic statistical computations and express the solution in an appropriate context.

The results supported the shift from a heavy emphasis on tests to the use of a project-based approach to not only teach the topics, but also show the students how the data analysis process is used to solve business problems. The course content was structured to include a strong emphasis on the teaching of Excel functions and formulas, data cleaning procedures, selection of appropriate data visualization techniques, and presentation of the results. Future assessment will be used to track the success of the new approach.

Effective use of Teaching Intervention and Simulation Games: A Study in a Supply Chain & Operations Management Course

Abstract

Dr. Chin-Yen Alice Liu [1]

1. *Texas A&M University – San Antonio*

According to a Gallup Poll, a whopping 92% of Chief Academic Officers believe that college graduates are prepared for the workforce, and students also tend to believe that they are ready to enter the workforce; however, only 11% of employers agreed and only a quarter of the supervisors described that the new hires are well prepared with critical thinking and analytic reasoning skills. The major discrepancy on the skills gap is particularly polemic when it comes to problem solving and decision-making. Although a critical thinker may not always be a good decision maker, a critical thinker will evaluate and analyze data before making decisions. Nowadays, employers highly seek candidates with strong decision-making skills not just for roles at a managerial level but also the entry-level of employees since everyone will face a challenging situation at some point and will need to act in the best interest of the stakeholders. Therefore, the education sector must acknowledge the discrepancy and adapt in line with this shift to reflect this essential and in-demand skills of the future. In other words, it is not just what you teach, but how you teach in the classroom. The purpose of this study is to assess the effectiveness of a teaching intervention designed to improve students' critical thinking skills and make a better operational decision through a simulation game in a Supply Chain & Operations Management course.

Has Higher Education Lost Its Way? Or Just Its Credibility?

Abstract

Dr. Hershey Friedman [1], ***Dr. Taiwo Amoo*** [1], ***Dr. Barbara Lewis*** [2]

1. *The koppleman school of Business,* ***2.*** *The Koppelman School of Business*

The authors posit seven fundamental reasons for higher education. They are: to procure skills for a career; to foster a love of learning; to develop critical thinking abilities; to acquire a philosophy of life and intellectual/cognitive stimulation; to pick up values, morals, and ethics and learn empathy; to become creative; and to learn tolerance and intellectual humility. This paper demonstrates how many colleges and universities have lost their way by offering courses that do not fulfill any of the above purposes. Unfortunately, many courses teach students *what* to think rather than *how* to think. Diversity of opinion has been replaced by indoctrination, and intellectual humility is superseded by intellectual arrogance. The best cure for the dangers of absolute certainty and overconfidence is humility and appreciating the value of doubt, constructive debate, and compromise.

Healthcare Analytics Teaching Tools: Cases for Ensuring Uniform Dates

Abstract

Dr. Yucheng Chen [1], ***Ms. Gwendolyn Powell*** [1], ***Dr. Chaza Abdul -Al*** [2], ***Dr. Loreen Powell*** [1]

1. *Commonwealth University of PA,* ***2.*** *Harrisburg University of Science and Technology*

The worldwide use and sharing of electronic healthcare data sets provides a rich source of information upon which can be analyzed for new discoveries for patient care, public health, and so much more. However, the sharing of world-wide health data sets also brings apparent issues when analyzing the data based on length of time or exposure. As a result, it is essential that healthcare analytic professionals understand the importance of ensuring uniform dates among their data sets. This research paper aims at providing current examples and cases in healthcare data sets where it is vital to have uniform dates among all data sets. Moreover, potential problems and issues from not providing uniform dates in healthcare data sets will be discussed. This research has practical impacts for health analytics practitioners, faculty, and future authors.

Inclusive Design Thinking: A Model for Architecting Inclusive STEM Classrooms for Women

Abstract

Ms. Pavani Sunku [1]***, Dr. Roderick Lee*** [1]

1. *Penn State Harrisburg*

Systemic inequities in student access and success have plagued higher education since its founding, especially for women The causes are complex and deeply rooted in the foundations of our society and educational institutions. Moreover, laws intended to transform higher education institutions such as the Morrill Land-Grant College Acts, the Servicemen's Readjustment Act of 1944 (G.I. Bill), and the Space Race had the effect of maintaining structural and systemic inequities. While higher education institutions have shifted their attention to addressing inequities in student outcomes, most interventions have not asked the students directly what they believe will contribute to more equitable outcomes. As a result, there is a need for a novel approach that centres on the needs of female students. In this research-in-progress, we leverage inclusive design thinking to create STEM classroom environments that support equitable outcomes for women students, especially those who have navigated higher education intuitions at a structural disadvantage. Inclusive design thinking includes multiple phases which often overlap. This process involves a conscious and deliberate effort of noticing, empathizing, and understand students and how they perceive their environments. The model then focuses on defining students' needs and identifying possible solutions that can be used to architect culturally inclusive and welcoming STEM classroom environments which can be quite advantageous for women. The last stage, testing, is based on feedback and a continuous improvement process. This concludes with implications for research and practice.

Keywords: Inclusive Design Thinking, Equity-Centered Design, STEM Education

Is Higher Education Becoming a Relic Because it Refuses to Adapt to Changing Conditions?

Abstract

Mr. James Lynch [1]***, Dr. Hershey Friedman*** [2]***, Prof. Chani Mintz*** [1]

1. *Brooklyn College,* ***2.*** *The Koppelman School of Business*

Whether higher education is in trouble is a hotly debated topic. Are higher education institutions in for a huge surprise or not? Should colleges and universities adapt quickly? One troublesome issue is the sharp drop in higher education enrollments as well as the decline in the number of colleges in the United States. The authors posit that colleges and universities must change if they will survive. There has been a disconnect between higher education and employability. College presidents have been more interested in creating administrative bloat (more schools, more academic departments, more assistant and associate provosts and deans) than in improving the quality of education. One recommendation is that schools must focus on providing marketable skills, including critical thinking, communication, and creative thinking. Given our tumultuous times, the most vital competency of all might be inculcating in students a passion for lifelong learning. This will enable them to use acquired skills to transform their companies or to find new jobs.

Keywords: higher education, organizational agility, value of college degree, critical thinking.

Nothing Cryptic About the Importance of Internal Control

Abstract

Dr. Jonathan Daigle [1], ***Dr. Ronny Daigle*** [2]

1. *Monmouth University,* ***2.*** *Sam Houston State University*

This instructional case explores the FTX saga, which led to its eventual bankruptcy. FTX has been heavily criticized for its lack of internal controls and questionable auditing of its financial statements. This situation provides a ripe opportunity to help students identify issues based on the Committee of Sponsoring Organizations of the Treadway Commission (COSO) framework, a widely recognized framework for internal controls and governance. Using the testimony of FTX's current CEO, John Ray, students are required to identify red flags with respect to FTX's control environment, risk assessment, control activities, information and communication, and monitoring.

Overcoming the Barriers to Participation in Quality High Impact Practices for Structurally Disadvantaged Students in STEM

Abstract

Dr. Roderick Lee [1], ***Mr. Adam Lenker*** [1]

1. *Penn State Harrisburg*

The Aspen Institute's College Excellence Program operationalizes student success along four key dimensions: learning, completion, transfer and bachelor's attainment, labor market, and equity in access and outcomes. Empirical research has demonstrated that high-impact practices (HIPs) substantially improve success in each of the four areas for all students, particularly among populations that have been historically underrepresented in STEM education (Kuh, 2008, 2016).

Despite the efficacy of HIPs on student success outcomes, students from structurally disadvantaged groups – such as Black, Latinx, and Native American, first-generation, and low-income – and numerically underrepresented students such as women in STEM are more likely to experience fewer HIPs. These students must contend with inequitable access to HIPs such as faculty-mentored research, study abroad, professional internships, honors programs, and living-learning communities that places them at a further disadvantage. Unfortunately, structurally disadvantaged students have noted barriers to access to HIPs such as lack of knowledge about HIPs, the benefits of HIPs, lack of time to figure out how to access and engage in HIPs, and prohibitive costs to participating in HIPs (Boyer 2023 Commission, 2022). To address these challenges, this research-in-progress articulates the barriers to accessing HIPs, actionable strategies to overcome the barriers, and strategies for leveraging technology to expand access to HIPs.

Preparing of Industry Ready Civil Engineering Students through Transformation of Capstone Design Course at UDC

Abstract

Prof. Pradeep K Behera [1]

1. *University of the District of Columbia*

Most of the engineering program curriculums include capstone design project courses in their senior year which include demonstration of knowledge and skills to solve one or more real world problems. Often this course is designed to accomplish several levels of achievements including knowledge, comprehension, application, analysis, synthesis and evaluation and they are ready to enter into professional practice.

The Civil Engineering curriculum at the University of the District of Columbia include two capstone design courses offered in the senior year. These courses have been comprehensively revised over last five years and the curriculum has been designed based on three considerations – (i) the real-world problems they are relevant to Washington DC Metropolitan area, (ii) ability of students to apply their knowledge and technical skills from courses previously taken from the curriculum, and (iii) effort to develop student's competence and confidence such that they can potentially become entrepreneurs. Over the two semesters, students go through the following sequence: (1) understanding of civil engineering systems through reading article and presentation, (2) small site development project to getting a permit for a client, (3) one large project such as development of a land of 30 acres for a client including layout of a sub-division and associated designs including stormwater management, storm sewer system design, sanitary sewer system design, water supply system design, multi-stories building structural engineering, transportation engineering, planning and scheduling, cost and sustainability. The paper will share the student outcomes, challenges, and opportunities from students and faculty instructor perspective.

Reflections on Teaching Information Visualization in Business Analytics Courses

Abstract

Dr. Peter Tarasewich [1]

1. *Northeastern University*

Information visualization is an essential part of business analytics (BA). It can be used as part of exploratory data analysis to look for potential relationships between variables, summarize the overall distribution of a data set, flag potential outliers, and other such activities to get a general feel for the data. Visualization can also be used to present insights, findings, messages, or stories found in a data set.

For our degree programs and courses in business analytics, we needed to decide how to best integrate the concepts and skills needed to understand good information visualization practices. Given the importance of the subject, we opted to require a complete course on information visualization as part of our graduate degree (MSBA) and for our undergraduate business major in BA. This course is also an option for our undergraduate minor (across all university degrees) and for specializations in our other business masters' programs (e.g., MBA). But we also needed to consider how information visualization fit into each of our other BA courses.

This presentation summarizes the information visualization topics that we currently cover in our BA courses, along with some of the teaching methods and activities that we have adopted. We are also preparing an introductory reading of information visualization topics meant to be used across our different courses, and hope to disseminate this as a teaching note for faculty at other institutions to use.

Shifting Forums : Establishing MS in Business Analytics Program

Abstract

Prof. Yang Lee [1], ***Dr. Peter Tarasewich*** [1], ***Dr. Bhawesh sah*** [1]

1. *Northeastern University*

This research examines how different forums initiates, and iteratively shapes institutional learning to establish an education program, namely the business analytics program. Specifically, the university administration, college administration, college faculty, university faculty, the external course design team, and the external industry advisory team worked together and separately at times to shape the discussion, and form the communities of forum, which orchestrated the establishment of the MS business analytics program. The timely emergence of a forum punctuated the pause and continuation of the elements of the program development initiative. We investigate the roles, decisions, and decisions of each forum in the context of shaping the local and global discussion of the program. At times, each forum exhibited its key interactions with other forums for revising the internal logic and without considering the external logic and forces. A few leading forums played a critical role for disrupting the local logic and the goals to reshape the discussions among the forums. The conflicts and collaborations among the forums are analyzed in the context of institutional learning. Key attributes for each forum's communication and knowledge strategies are also investigated. The research also discusses the limitation and implications for future research and practice. This research expands the institutional learning, knowledge management, and organizational learning theories in the context of inherent conflicting goals. In practice, faculty, administrators, industry observers, and curriculum or course design teams should benefit from this case of establishing the MS in Business Analytics Program in a university successfully.

Student Attitudes towards Multiple Attempts and Feedback on Quizzes

Abstract

Dr. Vitaly Brazhkin [1]

1. *University of West Florida*

It has been noted in academic literature that research on the use of multiple attempts for online assessments is sparse and inconclusive and lacks the voice of the students. To help bridge the gap, a survey of over 150 students across multiple supply chain management classes has been conducted. It has revealed that the students prefer three attempts on quantitative assessments in non-multiple choice formats. The preferences do not appear to depend on age, gender or GPA. However, there has been a statistically significant reduction in the number of the desired attempts compared to the data from the pre-COVID-19 period, which might be explained by students becoming more efficient learners during the peak of COVID-19 when all classes transitioned to the online format. Other findings indicate that students favor concrete feedback over abstract and would like to receive help with incorrect answers after their first attempt. Rather than have the correct answer given away, students prefer the type of feedback that allows them to solve the problem on their own. This study helps pave the way to better understanding of effectiveness of and student satisfaction with different assessment settings of online assessments of quantitative assignments.

The Impact of Live Lectures on Student Engagement and Performance in Online Classes

Abstract

Dr. John Weber [1], ***Dr. Bhupinder Sran*** [1]

1. *DeVry Univesity*

For many years, online courses have long been asynchronous activities. Such activities include online discussions, homework assignments, case studies, and research papers. When the COVID pandemic started, many universities began offering synchronous, live-lecture sessions within their online courses. At some universities, the live lecture requirement is optional, and at other universities, it is required.

This study will examine the relationship between student performance and live lecture participation in online statistics and business courses that offer a synchronous lecture component. Specifically, we will explore student grades, participation in other activities and overall time on task. We will compare differences in performance between those who attend the live lectures and those who remain purely asynchronous.

The integration of Collaborative University Business Experiences (CUBEs) and Collaborative Online International Learning (COIL) in Teaching

Abstract

Dr. Xiangrong Liu [1]***, Prof. Annett Grossmann*** [2]

1. *Bridgewater State University,* ***2.*** *HHN*

This teaching note shares the experiences of integrating Collaborative University Business Experiences (CUBEs) and Collaborative Online International Learning (COIL) together in teaching courses in Higher Education. This paper demonstrates the successful collaboration between a Service Operations Management course in Bridgewater State University, United States, and a Logistics & Business English course at Heilbronn University of Applied Sciences, Germany. Through working on a CUBEs project addressing the sustainability problems in the dining halls at the corresponding Universities, students in both classes worked closely through 6 weeks and benefited from the comparisons of sustainability practices between both universities. They provided strategical suggestions to both universities and made the final presentations for the representatives of the dining services. The survey data was collected in the end through Qualtrics and evaluated by the lecturers.

Visual Business Intelligence implementation in the Curriculum

Abstract

Dr. Anil Aggrawal [1], ***Dr. Cong Zhang*** [1]

1. *University of Baltimore*

As social networks diffuse so does the data generated through them. Data is generated at the speed of light and organizations must take advantage of it. Data have been defined to have four characteristics, four Vs(velocity, variety, volume, and veracity). Organizations must take advantage of this and other data that is available and can be explored both for middle and upper-level managers. Managers like charts that can easily be sliced and diced for problems and opportunities. Organizations are demanding graduates who can understand, manage and make sense of such data. Universities are obligated to provide such skills to their graduates. Developing such courses, however, is challenging since it requires knowledge from multiple disciplines such as information systems, social sciences, programming, etc.

Visualization is becoming an important tool for exploring data Visualization, typically, involves data exploration and explanation. Exploration may require programming, and presentation skills whereas explanation requires analyzing and interpretative skills. Interactive visualization provides slicing and dicing capabilities making decision-making in real-time a possibility.

A course may take many different forms depending on the audience. Business students may have soft skills and be better at analysis whereas technical students are maybe more savvy at exploration. Students' backgrounds can dictate the form of the course content and design. Many questions like, Does one course fits all? What are the controllable/uncontrollable variables? Skills needed to enroll in the course? Are platforms available? Etc. We are developing several courses related to visualization and data analytics.

We've got this: Creating High-Impact Online Capstone Experiences for Business Programs

Abstract

Dr. Vallari Chandna [1], ***Mr. Praneet Tiwari*** [1], ***Mr. Anup Nair*** [1]

1. *University of Wisconsin-Green Bay*

While the growth of online offerings in business programs across the world has been further accelerated by the pandemic, certain issues that typically plague online courses continue to weigh on instructors' minds. While online teaching is at times difficult but nevertheless interesting always, one of the challenges often faced by instructors is how to create an engrossing online capstone course for their business program? While there are no one-size-fits-all solutions, we endeavor to provide a brief overview on how instructors can utilize different processes and best practices to create a business capstone course. We discuss how "High-Impact" projects can be utilized to create a challenging yet engaging capstone course. Aspects such as typical challenges, mirroring of in-person classes, creating an instructor presence, and community building are discussed. Our aim is to share practices that have met with success with the aim of starting a conversation among B-school instructors, particularly those who are creating or updating their capstone courses so that they may come together and learn how to create greater student engagement.

Healthcare Analytics and Services Management

A CASE PRESENTATION OF A CLIENT WITH A DIAGNOSIS OF SCHIZOPHRENIA IN A COMMUNITY CARE UNIT IN MELBOURNE, AUSTRALIA

Abstract

Mr. Qian He [1]***, Dr. Michael Ha*** [2]***, Dr. Tan Kan Ku*** [3]

1. *Stockton University,* ***2.*** *Beijing Institute of Technology,* ***3.*** *Institute of Health and Management*

ABSTRACT

The Institute of Health and Management (IHM) in Melbourne, Australia, is aiming to deliver an Advanced Mental Health Nursing Unit at Australian Quality Framework (AQF) Level 8. This unit of study is a 12-weeks Speciality Stream focuses on teaching overseas undergraduate nurses (with a Bachelor degree in Nursing) to familiarise themselves with the Australian model of mental health nursing. This course will be delivered in 2021 as soon as it is being approved by the Tertiary Education Quality and Standards Agency (TESQA) in Australia. The focus is to teach international nurses the role of a case manager. A well written case study presented at a case conference will benefit the Multi-disciplinary Team (MDT) in formulating a rehabilitation plan for the client to adapt to the Recovery Model based on the consumer partnership paradigm.

INTRODUCTION

The Victoria's Mental Health Services provide a variety of treatment options for clients with special needs. The rehabilitation program is specifically related to residential services which provides community-oriented mental health initiative. Since 1995, several Community Care Units (CCUs) were built in various regional areas in Victoria (in residential setting) to replace the long-term rehabilitation wards of a psychiatric hospital. This case study aims to provide a model sample for international nurses to present at a Case Conference. This paper presents a case study of a 30-year-old single woman, Miss X (a pseudonym to ensure anonymity), diagnosed as suffering from paranoid schizophrenia, who was a resident at a Community Care Unit (CCU) in Melbourne.

A Case Study on Scheduling and Human Resource Allocation for Surgical Procedures

Abstract

Prof. Afrooz Moatari-Kazerouni [1], ***Dr. Amin Keramati*** [1]

1. *Assistant Professor, School of Business Adminstration, Widener University*

The performance of a healthcare organization relies on the availability and allocation of its employees. Health service providers commonly allocate their human resource by the type of labor functions. To improve this process, the human resource scheduling and allocation needs to be determined as a whole and be aligned with the mission, strategy, and the goals of the organizations.

This research performed a case study to analyze the processes of picking supplies and instruments for surgical operations at a teaching hospital. The studied hospital accommodated a new building, where some of its most critical care units were relocate. This improvement created opportunities for revisiting the processes among the hospital units, while new tasks and responsibilities were arisen for the hospital staff. The surgical process being a significant and one of the most influential processes in the hospitals, this study addressed the human resource allocation and schedule after the introduction of case carts as a new part of the surgical instruments and supplies picking process. For analyzing the imposed impacts of using the case cart, the new processes were mapped out and the activities' durations were estimated using the Maynard Operation Sequence Technique (MOST). The essential number of employees to perform instruments' collection for on time performance of surgeries is then estimated based on historical data on surgical schedules and the time-series forecasting.

Critical Access Hospitals and Their Financial Vulnerability

Abstract

Dr. Chamila Kalpani Dissanayake [1], ***Dr. Dinesh Pai*** [1]

1. *Pennsylvania State University*

We examine the changes in the performance of critical access hospitals (CAHs) and identified factors influencing financial performance. We use data on acute care hospitals in Pennsylvania's rural counties. Trend analysis and fixed-effects regression analysis are applied to the data set to investigate our objectives.

Effect of Internet Usage Behavior on the Influenza Vaccine Uptake Using Machine Learning Algorithms

Abstract

Mr. Rongxuan Wang [1], ***Ms. Chunjie Zhao*** [1], ***Dr. Yue Gao*** [1]

1. *Clark University*

People have been enjoying the dividends of Internet popularization, which has profoundly influenced people's health behaviors and their lives. This study was conducted to reveal the relationship between Internet usage and the vaccination status with machine learning algorithms. We collect information on Internet usage behavior related to health consultation, demographic, ethnic, educational, and occupational characteristics from the National Health Interview Survey (NHIS). Several common machine learning algorithms such as logistic regression, LASSO, random forest, and XGBoost will be implemented to predict the influenza vaccination decisions of respondents as well as to analyze the importance of relevant variables. Our results will facilitate policymakers to effectively leverage the significance of the Internet to promote prevalence of influenza vaccine.

Impact of COV-19 Pandemic on burnout and job satisfaction among emergency department health professionals

Abstract

Dr. Neset Hikmet [1]***, Prof. Menderes Tarcan*** [2]

1. *University of South Carolina,* ***2.*** *Eskisehir Osmangazi University*

Emergency medicine professionals (EMPs) are crucial members of the medical staff. Although they play an extraordinary role during the COVID-19 pandemic, they are mostly exposed to various health and safety risks that have significantly impacted their mental health, giving rise to symptoms, such as burnout and job satisfaction. This study investigated the relationship between the perceptions of COV-19 quality of life, burnout and job satisfaction of those working in two different hospital's emergency departments. A cross-sectional survey of two hundred and ninety-three participants was interviewed, using validated instruments (the Maslach Burnout Scale, the Minnesota Satisfaction Questionnaire and COV-19 Quality of Life Scale). Participants include 46 physicians, 105 nurses, 98 medical technicians, and 44 information technicians. Study findings indicate that significant relationship exists between quality of life and burnout and job satisfaction; annual income and household economic-well-being had a positive association with job satisfaction, whereas gender, age, education, marital status had no significant effect on any form of satisfaction. Also, the COVID-19 pandemic had a tremendous impact on the burnout levels of medical technicians. The findings of this study indicate that it is still unclear which variables significantly contribute to the relationship between burnout and job satisfaction. This research has the potential to provide light on the interconnected nature of these two problems, allowing health care managers to tackle them together rather than one at a time.

Keywords: Cov-19 Quality of Life,emergency medicine professionals, burnout, job satisfaction

ISO 9001:2015 Quality Standard Implementation: Constraints or a Source of Capabilities for Healthcare Resilience?

Abstract

Dr. Bill Ritchie [1]***, Dr. Steven Melnyk*** [2]

1. *James Madison University,* ***2.*** *Michigan State University*

Straining the resources of individuals, organizations, and governments, the COVID-19 pandemic revealed the importance of developing resilience capabilities and resources. This study aims to contribute to resilience research by exploring the role of sustainable process standards (e.g., the International Organization for Standardization (ISO 9001) family of quality standards) in bolstering organizational resilience amid crises and calamities. Drawing upon a case study methodology and survey data, this study endeavors to identify fruitful research directions related to the extent to which implementation of ISO 9001:2015 standards helped to build organizational resilience, as measured by the level of organizational preparedness for the disturbance.

Measuring Spectator Sports Quality: Data Analytics on Co-creation of Value

Abstract

Dr. Ahmet Ozkul [1], ***Dr. Gazi Duman*** [1]

1. *University of New Haven*

The article discusses the intersection between sport analytics and spectator sports service quality and focuses how IT and analytics can contribute to monitoring, analyzing, and predicting service quality in spectator sports from a holistic, co-creative multi actor point of view.

On the Role of Trust and Previous Psychological Contract Violation in Consumer Switching to Telemedicine Services

Abstract

Dr. Dodi Mossafer [1], ***Dr. David Gefen*** [2], ***Dr. Qizhi Dai*** [2]

1. *Accenture,* ***2.*** *Drexel University*

The use of telemedicine increased dramatically during covid-19 with providers, patients, and regulators favoring telemedicine over in-person visits. This study examines some of the reasons patients choose to switch to telemedicine, specifically "seeing a doctor" through mobile apps rather than in person. We draw on psychological contracts as adapted from consumers' perceived contract violation, aka psychological contract violation (PCV), and trust in a digital business environment. We argue that consumers consider switching from their current traditional providers to such online services based not only on their perceived value and trust in the provider, but also based on previous psychological contract violations they experienced with traditional providers. We conducted an online survey among individuals who had used healthcare services at least once in the past twelve months. Our data analysis results show that consumers were attracted to telemedicine when they had a high degree of trust in services they received from telemedicine providers. At the same time, consumers were pushed toward telemedicine by adverse perception of the traditional healthcare services they received before, specifically the perceived contract violation which reflects the discrepancy between the expected service level and the actual service experience. Moreover, consumers' switching intentions were facilitated by their perceived value of telemedicine. This study contributes to research on telemedicine adoption by investigating the factors that motivate consumers' decisions on switching to mobile apps. It also contributes to research on PCV by showing that PCV in offline experience can lead consumers to adopt the online service delivery mode.

Predicting Internet Use for Health Information: Evidence from the U.S. IPUMS Data

Abstract

Ms. Chunjie Zhao [1], ***Mr. Rongxuan Wang*** [1], ***Dr. Yue Gao*** [1]

1. *Clark University*

With the progressive growth of the Internet, E-health services have gradually penetrated people's lives, enhancing people's awareness of their own health attributed to their desire of higher living standards. This study investigates the factors (such as age, health status, sex, education, income, and occupation) that predicting the internet use for health information and services using machine learning techniques. Our analysis provides some implications on pattern of the internet use for health information across different demographic characteristics and shed lights on effective communication between patients and primary care providers.

Review of the Receiver Operating Characteristic (ROC) Curve Use for Medical Decision Making

Abstract

Dr. Yucheng Chen [1], ***Dr. Chaza Abdul -Al*** [2], ***Dr. Loreen Powell*** [1], ***Ms. Gwendolyn Powell*** [1]

1. *Commonwealth University of PA,* ***2.*** *Harrisburg University of Science and Technology*

The Receiver Operating Characteristic (ROC) Curve analytical test dates back to 1950s. However, its usage within the healthcare field began in the 1960 with diagnostic radiology on an existing detection of pulmonary tuberculosis data set. From there, the ROC Curve slowly began to be used in the healthcare research. Today, the ROC analytical test has become a preferred analytical test for accuracy of medical diagnostic systems because its accuracy is not contorted using cut-offs or specific thresholds. Instead, the area under the curve (AUC) derives a measure of sensitivity and specificity. While ROC Curve is preferred and valuable test for accuracy, it is often it is difficult to explain. To add to the literature and better explain the ROC Curve, this study provides a review of the ROC curve concept. Additionally, this paper provides a summary table of various clinical research studies that used the ROC Curve for medical decision making. This research has practical impacts for health analytics practitioners, faculty, students, and future authors.

Soft Skills as The Driving Force of the Fourth Industrial Revolution

Abstract

Dr. Haleh Karimi [1]

1. *Bellarmine Univeristy*

Nobel Prize-winning economist Heckman determined that having soft skills literacy statistically leads to success more than technical skills literacy. He cites evidence demonstrating that soft skills and competencies are *essential* for professional and personal success. Individuals with soft skills tend to have higher salaries than their counterparts. The Healthcare industry is on a rapid growth path, and professionals with a combination of soft and technical skills positively impact job promotion and wage increases, enabling improved individual performance and better organizational outcomes. The study interviewed twenty-seven healthcare employers to gauge the value healthcare employers place on demonstrating soft skills as we enter the fourth industrial revolution. Findings indicate that healthcare employers place a high value on soft skills competencies and recommend developing a successful pathway by investing in a sustainable soft skills educational system that trains students in interpersonal, professional, and leadership/management skills in their core areas. Given this research outcome targeting these behaviors is incredibly beneficial for healthcare professionals and organizations to manage humanity and emerging technologies to thrive within the fourth industrial revolution.

Keywords: healthcare, management, technology, soft skills

The Challenges of Healthcare Data Governance in Analytics

Abstract

Dr. Chaza Abdul -Al [1], Dr. Loreen Powell [2], Dr. Yucheng Chen [2], Ms. Gwendolyn Powell [2]

1. *Harrisburg University of Science and Technology,* ***2.*** *Commonwealth University of PA*

Today, almost every healthcare organization within the United States inputs, stores, and shares electronic data. As a result, electronic data is the foundation of almost every actional decision made within healthcare organizations. However, many haphazard data governance practices including faulty data, duplicate data, conflicting overlaid medical data, or missing data still exists. This research paper aims at studying the haphazard data governance practices and challenges of healthcare data governance in analytics. Moreover, some future IT related development questions regarding the impact of tying payment to performance, overwhelmed or frustrated end-users of the data systems, and frequent software updates are raised in this paper for further research. This research has practical impacts for health analytics practitioners, programs, faculty, and future authors.

The Effect of Socio-Demographical Characteristics on Leader Member Interaction During the Covid-19 Pandemic Process: An Evaluation of Healthcare Professionals

Abstract

Dr. Neset Hikmet [1]***, Prof. Gamze YORGANCIOĞLU TARCAN*** [2]***, Prof. Meltem SAYGILI*** [3]

1. *University of South Carolina,* ***2.*** *Hacettepe University,* ***3.*** *Kırıkkale University*

High-quality leader-member interactions are defined by intimacy and connecting with the leader, which inspires employees to contribute to the organization's and its stakeholders' success. Employees that cultivate a high-quality communicative relationship with their leader earn their leader's trust, allowing them to make difficult decisions with the assurance that their leader will comprehend their objectives. High-quality leader-member contact stimulates employees to contribute to the organization's success and address its shortcomings. This study aims to elucidate the opinions of employees in health care organizations about leader-member interaction, as well as the socio-demographic factors of health workers that influence these perceptions under the extreme conditions encountered during the Covid-19 pandemic. This research encompasses all health workers (physicians, nurses, health technicians, and other health workers) employed at a Turkish Medical Faculty Hospital.

What are Measurement-based Care (MBC) software service providers saying in a world where MBC can transform mental healthcare?

Abstract

Mr. Donald Jenkins [1]***, Dr. Josephine Namayanja*** [2]***, Ms. Rebecca Namubiru*** [3]

1. *Mirah,* ***2.*** *University of Massachusetts, Boston,* ***3.*** *Hochschule Neu-Ulm*

Measurement-based care (MBC) is the systematic collection of patient-reported measures to monitor treatment progress and inform clinical decision-making. Past research shows that MBC can effectively improve patient engagement and healthcare outcomes, including in behavioral healthcare (BHC). Despite extensive developments of MBC in physical healthcare within the United States, research shows that only 11% of psychologists and 18% of psychiatrists apply MBC in their therapy. This slow uptake of MBC in BHC motivates our work. There is a growing base of software service providers offering MBC products and services to mental health professionals today. To better understand how these service providers are trying to shape MBC adoption within BHC, we examined themes from the MBC providers' perspective. Specifically, we captured 1,719 documents from 22 service provider websites, then applied the Latent Dirichlet Allocation (LDA) topic modeling method to identify language patterns and related topics within these documents. To identify the best-fitting model, we assessed models quantitatively, with coherence metrics, and qualitatively, with the assistance of clinically trained subject matter experts. Initial results identified 15 different topics we are now further analyzing for competitive analysis of the MBC software and services marketplace. This competitive analysis includes identifying distinct clusters of companies with common dominant themes, searching for patterns across different vendor types, gaps in the topics covered by specific vendors, and correlation between topics indicating potential hidden relationships in the documents. Further, the LDA model can be applied to future documents posted by MBC providers to monitor their dominant themes over time.

Human-Technology Interface

The Design and Operation of Digital Platforms under Folk Theories of Sociotechnical Systems

Abstract

Prof. Jordan W. Suchow [1]

1. *School of Business, Stevens Institute of Technology*

We consider the problem of how a platform designer, owner, or operator can improve the design and operation of a digital platform by leveraging a computational cognitive model that represents users' folk theories about the platform as a sociotechnical system. We do so in the context of Reddit, a social-media platform whose owners and administrators make extensive use of shadowbanning, a non-transparent content moderation mechanism that filters a user's posts and comments so that they cannot be seen by fellow community members or the public. After demonstrating that the design and operation of Reddit have led to an abundance of spurious first-party suspicions of shadowbanning in cases where the mechanism was not in fact invoked, we develop a computational cognitive model of users' folk theories about the antecedents and consequences of shadowbanning that predicts when users will attribute their on-platform observations to a shadowban. The model is then used to evaluate the capacity of interventions available to a platform designer, owner, and operator to reduce the incidence of these false suspicions. We conclude by considering the implications of this approach for the design and operation of digital platforms at large.

Innovation and Creativity

Neighborhood Food Sovereignty as Key Ingredient for Urban Innovations and Equitable Development

Abstract

Dr. Samia Kirchner [1]

1. *Morgan State University*

This proposed paper provides an alternative to the problematic development concepts in city development that perpetuate neighborhood (re)segregation as urban areas are gentrified. It focuses on the city of Baltimore, where 15,000 vacant lots and properties are being turned into lucrative real estate developments. A comparative analysis focused on food consumed at two public schools in two distinct neighborhoods, one where residents earn lower than median income of the State of Maryland and the other with highest, reveals the co-relationship between children's access to healthy and nutritious food access and test scores in the areas of art and math, the two subjects that form the basis for innovative and creative thought. The proposed paper underscores neighborhood food sovereignty approaches as critical to eliminating urban inequality, environmental and health crises. Food sovereignty is a food ecosystem where the consumers have access to and control the entire cycle. Developing urban farms in marginalized communities, where majority of the vacant properties are located, will allow school children to engage in the agroecological science while enhancing their ability to think creatively.

Programming spaces that support neighborhood food sovereignty is key ingredient for developing mental autonomy, innovation, creativity and comprehensive wellbeing of its residents.

Smart Farming for Sustainable Agriculture

Abstract

Prof. Edward Chen [1]

1. *University of Massachusetts Lowell*

The farming industry has been around since the beginning of civilization. For a while, the process of operations has been done the same way. The world of technology has adapted and shifted into many sectors of business and most recently the biggest shift in the agriculture business is smart farming. The demand for agricultural products continues to grow in many fast-growing metropolitan areas. The productivity and efficiency of farming need to be innovated to support the rising population. Smart Farming has been emerging as a form of transforming traditional agriculture. The use of smart farming is designed to optimize the overall farming experience, by using new technologies that allow corporations to analyze data to ensure maximization of productivity is being reached without giving up the quality of goods. Companies are switching to smart farming techniques to ensure they are reaching higher output levels resulting in increased profits. In this research, smart farming supported by information technologies such as Artificial Intelligence (AI) and Internet-of-thing (IoT) will be introduced to enhance sustainable agriculture.

Technology Innovation for Small Businesses – Creating an online presence during a time of rapid change

Abstract

Dr. Jas Gaurav Singh [1], ***Dr. Elif Kongar*** [2]

1. *Nichols College,* ***2.*** *University of New Haven*

The Covid-19 pandemic of the last four years has significantly disrupted the business landscape. Reductions in economic activity for several consecutive quarters affected small businesses more as they traditionally have less reserves compared to larger-sized businesses. This changing dynamic forced small businesses to innovate rapidly and utilize technology more effectively. As a result, trying to find additional revenue streams, small businesses, especially the ones that are heavily reliant on foot traffic, started seeking ways to reach their audience and engage their potential customers digitally.

One of the first steps that small businesses had to take was either to create a new online presence or to update their online presence for more effective interaction with their consumer segments. This required rethinking their business processes making sure that the online presence platforms reflected their conventional offline presence. This study outlines the steps various small businesses undertook to create an online presence for consumer facing operations. The impact of the technological improvements on this evolution is also discussed. The plug and play features in marketing, sales, payments, servicing, and cross selling are detailed. Operationalizing this new platform in a resource-scarce low-capital environment is also included in the study.

The overarching goal of the study is to provide small business a perspective on the need for constant innovation. Providing a roadmap for such scenarios, the success factors required by small businesses are explained.

Keywords: Innovation, Small-business, Technology

Legal, Ethical, and Social Issues

Built-In Bias: The Ethical Implications of Bias in Artificial Intelligence

Abstract

Dr. Rhoda Joseph [1]***, Dr. Mohammad Ali*** [1]

1. *Penn State Harrisburg*

Abstract:

Artificial Intelligence (AI) has seen rapid growth and interest in multiple disciplines. The primary advantage of engagement with AI is the increased ability of computer-based systems to interact with more human-like precision (West, 2018). However, human-like engagement through AI has been presented in popular culture as both assistive and destructive (Banks, 2018). As machines become more human, the biases that exist in human cognition and behavior are also embedded and coded into AI solutions.

Bias in AI is a concern that has been documented in several different scenarios including bias in AI systems used in criminal justice systems (Upadhyay & Romashkin, 2023) and bias in financial services (Bajracharya et. al, 2023). The prevalence of AI in all aspects of society including highly sensitive areas such as healthcare and security has increased ethical concerns about the governance of such systems (Agbese, et. al, 2021). Bias exists in many forms and one study proposed a three-pronged approach to addressing AI bias: understanding the bias, mitigating the bias, and accounting for the bias (Ntoutsi, et.al, 2020).

The primary research questions governing this study ask: what are the ethical considerations when building AI solutions with embedded biases, and how do end-users/consumers perceive bias and ethics in AI? This study will proceed along two dimensions. First, we will conduct a meta-analysis of studies that examine bias and ethics in AI, and second, capture primary data from end-user/consumers to understand their perceptions about bias and ethics in AI in the area of social media usage.

Reciprocality of Relationships Helps Increase Resilience: Strengthening Social Support through Social-Behavior Ceuticals

Abstract

Dr. Neset Hikmet [1], ***Dr. Hatice ATILGAN*** [2]

1. *University of South Carolina,* ***2.*** *Eskisehir Osmangazi University*

A vast majority of research confirms that social support for vulnerable (such as old, sick, injured, and handicapped) people may serve as a buffer against the impact of low levels of psychological resilience on mental health during crises. However, when the time of a crisis extends, providing social support becomes a burden for familial, official, and/or voluntary providers as well as receiving social support makes receivers often feel dependent and worthless. People receiving social support often think they have lost most of their mutually beneficial social relationships, which establish social connections and self-worth. Prior research shows that reciprocity in relationships has profound impacts on the emergence of social cohesion and self-worth. Understanding the microstructural changes in the one-to-one relationships of people in need and their supporters and implementing social-behavior interventions produced from social scientific theories may provide positive changes in both parties' lives. Building upon previous research, we claim that inserting active roles and reciprocal relationships into existing social support systems may provide a more effective and sustainable approach.

These behavioral interventions are effective and low-cost prescriptions without side effects in terms of increasing the resistance of individuals against social and psychological problems caused or accelerated by a crisis, keeping their morale high in the process of fighting the difficulties during a crisis, and contributing to their general well-being after a crisis. Hereby, individuals can better handle personal (health problems, aging, handicapping, etc.) and broader (e.g., pandemics, disasters, etc.) levels of crises.

Reframing and Restructuring Organizational Strategies for Addressing Harassment and Discrimination in the Workplace

Abstract

Dr. Erin Percival Carter [1]***, Dr. William Obenauer*** [1]

1. *University of Maine*

Harassment and discrimination in the workplace can have a negative impact on the well-being of employees and the overall performance and success of an organization. Left unaddressed, these behaviors can create a toxic work environment and lead to high levels of turnover, absenteeism, and decreased productivity as well as legal liability. Yet, many employees who experience harassment and discrimination are hesitant to report due to oft-justified concerns about retaliation. In this paper, we offer three suggestions for how organizations can better position themselves to address institutional shortcomings that may perpetuate harassment and discrimination. First, reports of harassment and discrimination should be re-framed not as threats to the institution to be managed but as reporting individuals providing a valuable service by raising awareness about vulnerabilities in existing systems akin to "bug bounty" programs common in the technology industry. Second, work to identify situations most likely to give rise to harassment and discrimination by considering how power shapes thinking and behavior in ways that can increase the likelihood of harassing or discriminatory behavior. Finally, seek not only to identify interventions likely to mitigate these risks but to ensure that interventions are applied at the optimal time to prove effective.

Operations Management / Operations Research

A Fuzzy Linear Physical Programming Approach for Performance Evaluation and Benchmarking

Abstract

Dr. Gazi Duman [1], ***Dr. Elif Kongar*** [1], ***Prof. Surendra M. Gupta*** [2]

1. *University of New Haven,* ***2.*** *Northeastern University*

The literature offers a wide range of performance evaluation methodologies to assess the efficiency of similar business entities. However, the majority of these approaches use weighted evaluation criteria to introduce the relative importance of each weight into the problem environment. These weights are rarely generated by a data-driven process which can lead to subjective results. This study presents a linear physical programming-based performance evaluation approach that eliminates the subjectivity in weight assignment. A case study is provided demonstrating the applicability of the approach.

A new data-driven multivariate process capability index

Abstract

Dr. SANGAHN KIM [1], ***Dr. Mehmet Turkoz*** [2]

1. *Siena College,* ***2.*** *William Paterson University*

Nowadays, organizations and companies continuously strive to achieve high levels of production and quality performance. In the modern industry, a process is typically monitored and controlled in statistical ways. Along with the philosophy of six sigma, a process is measured by a single index to determine whether a process is capable of producing products and services within the specification, i.e., capable of a certain quality performance presented by producers or customers, so called process capability indices (PCIs). So far, several univariate PCIs and only a few multivariate PCIs have been introduced. Due to the complex nature of multivariate processes, the methods used to assume specific well-known distributions such as a multivariate normal distribution. This work develops a new data-driven multivariate PCI using a support vector data description (SVDD) without the assumption of normality or any known probability distribution for the process data. It is significant in that no research has been carried under the distribution-free environment, to the best of our knowledge. It is also expected that the data-driven PCI will provide a practical usage in the actual processes since it does not depend on any specific distributions of the process data.

A Study of Agile and Lean Supply Chain Design

Abstract

Prof. Gang Li [1], ***Prof. Yusen Xia*** [2]

1. *Bentley University,* ***2.*** *Georgia State University*

Should a firm build a longer or shorter supply chain? A shorter supply chain seems an obvious answer. One can argue that by "cutting the middle-man", a company is able to respond consumer demand faster meanwhile saving cost. We build a simple model to demonstrate a counter intuitive result: Under some conditions, having a longer supply chain can save money and respond to the consumer faster. In this talk, we discuss how a company can design an agile and lean supply chain that provides consumers guaranteed service time with minimum cost.

Airline Ancillary Revenue Willingness to Pay during Covid 19 Pandemic

Abstract

Dr. Fouad Mirzaei [1]

1. *University of North Texas*

Airlines are facing challenges of intense competition and reduced revenues amidst the COVID 19 pandemic. This pressure has made the industry universally look for opportunities to generate ancillary revenue as a supplemental source of revenue. Measuring airline travelers' willingness to pay is crucial in pricing and estimating ancillary revenue demand. In this study, we used the theory of planned behavior (TBP) to understand users' willingness to pay for ancillary fees and assess how risk perception affects users' willingness to pay.

Cross-Channel Policies in Omnichannel Operations

Abstract

Mr. Tao Xu [1], ***Dr. Wenjing Shen*** [1]

1. *Drexel University*

Omnichannel retailers provide several cross-channel policies such as buy-online-pick-up-in-store(BOPS) and buy-online-return-in-store(BORS) to help customers take advantage of the omnichannel services. BOPS policy allows customers to place online orders and use an in-store pick-up option while BORS policy allows customers to buy the products online and return unsatisfied items in store. Under these policies, customers enjoy more flexibility and as a result, retailers attract more demand. On the other hand, these policies also create more challenges in omnichannel retailers' operation decisions compare to the conventional retailers since they integrate both online and offline channels. In this paper, we investigate how such policies affect an omnichannel retailer's product location decision, i.e, which product should be carried in the physical store, and the inventory decision, i.e, how much inventory should be carried. We also study how these results change under capacity constraint and propose a heuristic algorithm. We find that retailer's product location decision and inventory decision are affected by BORS and BOPS policies in the following ways: (i)BORS only: physical stores should store less inventory. (ii)BOPS only: physical stores should always store more inventory to satisfy additional demand from pick up orders. (iii)Applying both policies: physical stores should adjust their inventory level based on the fraction of high valuation customers. We also show that with higher return handling cost for returned products, BORS and BOPS policies lead to more store SKUs.

New model for measuring criticality in project management networks

Abstract

Dr. Fariborz Partovi [1]

1. *Drexel University*

This paper is about identifying the most important (critical) activity in a project management network. It indicates which activities require the most management attention. In CPM networks the number of critical paths along which an activity lies forms the importance of the activity. Unfortunately, in the PERT network because of the probability nature of activities the problems are much more complicated. In this paper, we will introduce two new methods based on social networks for prioritizing important activities in a network.

Optimal Coal Mine Allocation and Delivery Schedule for a Power Company

Abstract

Dr. nasreddine saadouli [1]

1. *Gulf University for Science and Technology*

A Power company uses high moisture and ash with low calorific content coal. The company needs to improve upon the present grade of coal being used through a network of coal suppliers from all over the world within time and budget constraints. A linear programming formulation is used to determine the optimal supply of coal from different mines in various countries depending on a number of technical and chemical specifications of the coal. Consequently, and to co-ordinate all activities, a delivery schedule is developed and monitored using the critical path method. A case study of a real company is studied, and the results are analyzed and discussed.

Optimum Placement of Electric Scooters to Pickup Stations

Abstract

Dr. Rajeev Kumar [1]

1. *Kutztown University of Pennsylvania*

Shared mobility systems including Bike Share schemes and Electric scooter (e-scooter) rental companies (such as Bird, Lime, and Uber) have proliferated in cities worldwide. The literature on Bike Share schemes has proposed methods that utilize users (riders) journeys data and other data (such as weather, day of the week, etc.) to estimate Bike Share demand. The proposed approach of this paper assumes that the daily demand of e-scooters is first estimated for different neighborhoods of a city by utilizing the existing approaches. The approach then utilizes estimated daily demand data in an integer programming model to calculate the optimum quantities of e-scooters at the pickup locations. The objective of the integer programming model is to minimize the distance travelled by the potential riders to the e-scooters. This approach can improve the utilization of the scooters, and thus, increase the profit of the scooter rental company and its fleet managers. The benefit of the proposed approach over the existing approach is discussed in the paper.

Properties of job sequences for the three-machine flowshop problem with sequence-dependent family setups to minimize makespan

Abstract

Prof. Jeffrey Schaller [1]***, Dr. Jatinder Gupta*** [2]

1. *Eastern Connecticut State University,* ***2.*** *University of Alabama Huntsville*

This paper considers the three-machine permutation flow shop problem to minimize makespan with sequence-dependent family setups. The group technology assumption is used, in which once a family starts processing on a machine, all the jobs are that family are processed before another family starts processing on the machine. In this problem, both a sequence of families as well as a sequence of jobs within each family is required to define a solution. This presentation will provide some properties that can be used to determine if a job sequence within a family is optimal. The presentation will also provide some avenues for using these properties.

Understanding Opaque Selling from an Inventory-control Perspective

Abstract

Mr. Yuan Qu [1], ***Dr. Jian Yang*** [1]

1. *Rutgers University*

We take a mainly inventory-control angle in our attempt to understand the roles played by opaque selling. While facing demands for two closely-related products, the seller in our model can also use price discounts and more guaranteed availabilities to entice buyers into accepting probabilistically assigned actual deliveries. Our study involving such opaque selling first takes aggregate buyer behaviors as given. Here, the seller's provable preference for more balanced inventory levels would in turn lead to both a very reasonable balance-inducing rationing policy and the seller's welcome of more opaque selling. Different from existing works with similar purposes, we allow the seller's replenishment policy to tolerate occasional lost sales. It would be better able to extract operational efficiencies in absolute terms and rationalize opaque selling by comparison. Also considered is a more realistic but also more complex case where the choices of a continuum of buyers with heterogeneous valuations and the fulfillment rates of the lone seller are reached as equilibrium points of a game-theoretic model. The seller's decisions on the price discount and replenishment policy are discussed as well. Further managerial insights are gained from our numerical analyses.

Sustainability Management

Are Indoor Vertical Farms Sustainable?

Abstract

Dr. Eric W. Stein [1]

1. *Penn State*

Billions of dollars in venture capital has been put into indoor farming; i.e., growing leafy greens and other vegetables in indoor controlled environments. The question is: are these ventures sustainable from a triple bottom line perspective?

The author will summarize research on the profitability of indoor farms, their impacts on air, soil and water, and the energy requirements of these systems. The author will also present the creation of an online assessment tool that calculates operational metrics for indoor farms. Key measures include water use, labor, energy, capital, and space. The presentation will close with ideas for research on indoor farm design, layout, and optimization to decrease overall costs and maximize production. Attendees are encouraged to engage in discussion on ways indoor farms could be made more sustainable.

Corporate Governance and Voluntary Disclosure: The Moderating Role of Industry

Abstract

Dr. Nabil Tamimi [1], ***Dr. Rose Sebastianelli*** [1]

1. *University of Scranton*

Corporations face increasing pressure from stakeholders, including investors, to voluntarily disclose information on environmental and social policies, practices, and performance. ESG (Environments - Social - Governance) scores serve as an important signal to stakeholders of a firm's commitment to transparency and accountability. Our study focuses on the S&P500. We propose and test a structural equation model (SEM) that considers the effects of corporate governance (diversity, autonomy, tenure, size) on the voluntary disclosure of environmental information (E score) and how this impacts the voluntary disclosure of social information (S score).

Based on data retrieved from Bloomberg, we find board diversity, board autonomy, and board size to positively impact E-Score; board tenure negatively affects E-Score. All path coefficients from corporate governance characteristics to environmental disclosure are significant. We also find the path coefficient from E-Score to S-Score to be strong, positive, and significant. This indicates that environmental disclosure mediates the relationship between corporate governance and social disclosure.

Industry sectors perceived as more polluting (e.g., utilities) are under even more scrutiny to be transparent. Consequently, the model was estimated separately for two groups of companies, those belonging to sectors considered polluting and those belonging to sectors considered non-polluting. The estimated path coefficient for E-Score -> S-Score was found to be significantly higher for companies in polluting industries compared to those in non-polluting industries. This suggests that industry sector moderates the relationship between the voluntary disclosure of environmental and social information, with the relationship being significantly stronger for those companies in polluting industry sectors.

Price Gouging at the Pumpkin Patch? Expense Neglect in Agritourism Leads to Perceptions of Price Unfairness

Abstract

Dr. Erin Percival Carter [1]

1. *University of Maine*

People often bemoan the seemingly high prices associated with seasonal agritourism activities such as picking berries, apples, pumpkins, or Christmas trees. From the consumer's perspective, it can seem unfair to have the prices for agricultural products the consumer harvests equal or even exceed the market rate for similar products in the grocery store which presumably incurred additional harvesting and transport expenses. This perspective, however, fails to account for the unique expenses and other challenges to profitability associated with engaging in agritourism from a farmer or other agricultural producer perspective. In this paper, we use behavioral experiments to examine the phenomenon of expense neglect, how it can lead to perceptions of unfair prices, and how highlighting the unique costs and challenges of allowing public access to agricultural activities can raise the cost of production for agricultural goods and thus result in increased willingness to pay and the perception that prices are more fair. We conclude by discussing implications not only for agritourists and agritourism providers but for government and advisory groups considering suggesting agritourism as a potential solution for small and medium farms attempting to reach profitability.

The developmental path of firm sustainability in green decision and transformation

Abstract

Dr. Tung-Shan Liao [1], ***Mr. Jia-Hang Ruei*** [1]

1. *Management College, Yuan Ze University*

Through examining the shift of business paradigms brought by ESG (environmental, social, governance), it has driven many industrial organizations to try to respond through the green transformation. Still, there is a need for an in-depth discussion on how firms can transform from the so-called financial or profit model to the model in which firms tend to benefit the shareholders thru the ESG approach. This study explores how firms conduct green decisions and transformations to respond to green forces. By using the grounded theory with 9 cases (all of them are listed companies in Taiwan), we analyze their operations, strategies, developmental paths, and the mindsets of the high-level managers for their responses to the green issues. The findings of this study point out that in the face of corporate sustainability (CS), firms build their capabilities in green transformations through the ambidexterity innovation mode. Either actively or passively responding to the green forces, firms reinforce their sustainable influences through dialogue with stakeholders. In addition, they create their green competitiveness through shared value creation. This study also finds that firms' CS practices follow the driving dimension (by the green forces), the action dimension (through innovation, dialogues, and value creation), and the effect dimension (through green operation and the co-creation of shared sustainable values), for the responses of ESG. This study proposes a "three-dimension developmental model" for sustainable corporate transformation based on these findings. By this model, we clarify how firms build their green transformation capabilities and respond to green trends.

The relationship between logistics performance and environmental quality

Abstract

Dr. Youqin Pan [1]

1. *Salem state university*

Logistics industry is critical to a country's economic development. However, it can also play an important role in environmental quality since it is considered as a major energy-consuming sector. In the current literature, there is a lack of research studies which investigating the relationship between logistics performance and environmental quality. The findings reveal that certain components of Logistics Performance Index significantly affect the quality of environment in countries selected for the study.

Strategy and Organizational Behvior

A Basic Leadership Inquiry – Failure of Leadership at Wells Fargo: A Case Study

Abstract

Dr. Sergey Ivanov [1], ***Ms. Cindy Hernandez*** [1]

1. *University of the District of Columbia*

Wells Fargo & Company also known as Wells Fargo is a multinational depository financial institution that was established in the year 1852, to provide banking services to the public. Wells Fargo believes that their customers are better served when they have a relationship with a trusted provider bank that knows them well, provides reliable guidance, and can serve their full range of financial needs, thus their mission statement is "helping customers succeed financially."

Wells Fargo prides itself by being one of the top three leading financial service institutions. According to "20 Largest Banks in the U.S." by Ruth Sarreal, Wells Fargo sits in the top three banks in the United States right next to Chase bank and Bank of America. They currently own approximately 1.69 trillion in assets, serve one out three households and more than 10% of small businesses. They have approximately 4,700 branches all over the country and employed about 268, 531 individuals.

Although Wells Fargo is one of the oldest pristine financial institutions in the United States and they are proud to share their mission, "helping customers succeed financially", they have not always done the right thing for their customers and their employees. In October of 2016, they were involved in a scandal in which the company created many fake accounts which included deposit and credit products for customers without their knowledge and consent.

A Study on the Effect of Internal Open Innovation Activities on Digital Transformation Capability and Firm Performance: Evidence from Taiwanese SMEs

Abstract

Prof. Chieh-Chih Hou [1]

1. *Ming Chuan University, MCU*

This study focuses on whether external knowledge inflow activities under open innovation, including financial knowledge acquisition, nonfinancial knowledge exploration, and strategic alliances, have a positive impact on firms' execution of digital transformation and whether firms' execution of digital transformation contributes to firm performance. In addition, we assess whether the knowledge absorption capability of enterprises interferes with their digital transformation capability. In this study, we collected 472 valid samples from Taiwanese SMEs, and we conducted data analysis using SPSS25.0 and Smart PLS3.0. The results show that all three types of external knowledge inflow activities have a positive impact on implementing digital transformation, digital transformation can effectively improve the performance of enterprises, and different interference effects occur according to different knowledge inflow activities.

Diversity and Inclusion Matter: We Are Better Together

Abstract

Dr. Sergey Ivanov [1], ***Ms. Bonita Harrison*** [1]

1. *University of the District of Columbia*

Abstract. Diversity and inclusion are important elements for all organizations to succeed. This paper will expound on why diversity and inclusion matter. It will also explain the differences between diversity and inclusion. Additionally, the paper will provide benefits, as well as barriers to a diversified workplace.
Keywords: Diversity, Inclusion, Workplace

Effect of External Financing on Sustainable Innovation: The Threshold Effect

Abstract

Dr. Xun Zhang [1]***, Dr. Biao Xu*** [2]***, Dr. Ning Zhang*** [3]***, Dr. Jun Wu*** [4]

1. *Hohai Univercity,* ***2.*** *Nanjing University,* ***3.*** *Northeastern University,* ***4.*** *Georgia Gwinnett College*

Considering 12,325 green enterprises in China between 2011 and 2013 as sample, this paper investigates the threshold effect of firm debts and government R&D subsidies on green enterprise's innovation, and further explores the effects and differences regarding high tech enterprises and enterprise's position in industry. The results show that, first, firm debts and government R&D subsidies have a positive threshold effect on green enterprise's innovation. When the leverage is less and government subsidies are greater than the threshold value, the contribution of firm debts and government R&D subsidies to green enterprise's innovation is significantly enhanced. Second, when the government subsidies are greater than the threshold value, the effect of government R&D subsidies to sustainable innovation for high-tech enterprises and industry followers is significant. Third, the promotion effect of debts on sustainable innovation for non-high-tech enterprises and enterprises with different position in industry is significant enhanced when the leverage is smaller than the certain level. On the basis of these result, enterprises can provide support for innovation activities through appropriate debts when the firm debts less than the threshold. What's more, the Chinese government should focus on strengthening subsidies for high-tech enterprises to achieve the optimal level of subsidies utilization efficiency.

Stakeholders, CSR, and Sustainability: A comparison of firms in extractive and non-extractive industries

Abstract

Prof. Shobha Das [1]

1. *IFMR Graduate School of Business, Krea University*

Extractive industries produce essential inputs for the global economy, but they also have a major effect on environmental sustainability. Stakeholders try to influence firms in these industries regarding their environmental impact. Is there a difference in the influence on, and CSR actions of, firms in extractive and non-extractive industries? This paper statistically analyzes primary data collected from firms in Qatar, a country that is reliant on oil and gas, an extractive industry. The results show that while the government is an important stakeholder for firms in both types of industries, the influence of customers and suppliers is significantly different. Similarly, the environmental CSR actions pursued by firms is different in the two types of industries.

The Dynamics of Paying it Forward: How and Why Leaders Transfer Past Generosity To The Future

Abstract

Dr. long wang [1]

1. *City University of Hong Kong*

We extended both the imprinting theory and reciprocity literature to investigate the dynamics of generosity transmission. Three complementary experimental studies based on behavioral economics games consistently showed that the prior receipt of generosity led leaders to pay forward generous compensation to their own employees after the leaders had already reciprocated the generosity they received before in kind. In addition, the governing structure and quality of exchanges between leaders and employees moderated how leaders paid forward generosity. Specifically, Studies 1 and 2 showed that when leaders were generously treated before, they were more likely to pay forward generosity to their employees as a result of their prior experiences. However, the generosity leaders paid forward only persisted when employees had the opportunity to pay back the leaders' generosity. Study 3 showed that leaders who were treated generously before were more likely to pay the generosity forward to their own employees when their employees reciprocated the leaders' generosity with high effort. Thus, the generosity leaders paid forward was intensified by a healthy cycle of positive reciprocity between leaders and employees. Our research contributes to the organizational research on imprinting theory and reciprocity by enhancing our understanding of how leaders carry forward generosity in economic exchange relationships.

Veteran Entrepreneurship: Review of Education Programs, Theory and Practice

Abstract

Dr. Sergey Ivanov [1], ***Mr. Nawras Taffal*** [1], ***Dr. Alex Maritz*** [2], ***Dr. Julius Anyu*** [1]

1. *University of the District of Columbia,* ***2.*** *La Trobe University*

Abstract. Entrepreneurship education programs have experienced significant growth and development globally, with the U.S. often leading the transformation pathway, particularly in areas of minority entrepreneurship, which include veteran entrepreneurship education programs supported or encouraged by the U.S. government. Our aim is to explore and investigate a first of its kind study on the status, distribution, and impact of veteran entrepreneurship education programs across the U.S., with the purpose to inform the body of knowledge on interventions to accelerate veteran self-employment prevalence. We do so by applying replication and extension research using emergent enquiry and muti—method research design and analysis, incorporating an entrepreneurial ecosystem approach that considers interactions and interdependencies between dynamic program components.

Keywords: Entrepreneurship, Entrepreneurship Education Programs, Veterans, Entrepreneurship Theory

Supply Chain Management and Logistics

Blockchain-based lithium-ion battery closed-loop supply chain safety passport

Abstract

Ms. Zhuowen Chen [1], ***Prof. Joseph Sarkis*** [1], ***Dr. Abdullah YILDIZBASI*** [1], ***Prof. Yan Wang*** [1]

1. *Worcester Polytechnic Institute*

Lithium-ion batteries (LIBs) are being widely used in electric vehicles and other products. Closing the loop of the lithium-ion battery (LIB) supply chain by reusing, remanufacturing, and recycling spent LIBs is an important solution to the metal resource constraints, environmental damage, and social impacts caused by the usage of LIBs. Safety issues are a main cause of disruption in the LIB closed-loop supply chain (CLSC). The disclosure, authenticity, and security of safety-related information, including battery information, transaction information, and operation information, are difficult to guarantee. This paper proposes a contribution to apply blockchain (BC) technology to manage information in the CLSC of LIBs to mitigate safety issues. The security, decentralization, traceability, investigation, and smart contract features of BC technology are utilized to manage the safety-related information of various LIB CLSC processes in BC. BC application in supply chain is still in early progress, we propose future implications and directions for the development of BC technology in CLSC.

Collaborative Logistic Vehicle Routing Problem Among Small Shippers: A New Genetic Algorithm Approach

Abstract

Dr. Vahid Ghomi [1], ***Dr. Sina Shokoohyar*** [2], ***Dr. Farnaz Ghazi Nezami*** [3], ***Dr. Vahid Dardashti*** [4]

1. *Penn State Mont Alto,* ***2.*** *Seton Hall University,* ***3.*** *Kettering University,* ***4.*** *Georgia Institute of Technology*

This is an extended version of VRPSD to solve collaborative logistics vehicle routing problems with split deliveries (CoVRPSD). The origin of each vehicle route can be one of the suppliers. We assume that the total vehicle capacity in the supply chain is greater than or equal to the total demand. However, the total capacity of each supplier might not cover its own demand, and therefore, we implement collaboration in the supply chain. We performed a series of numerical experiments and utilized a genetic algorithm to determine the position of the proposed model in comparison with the other VRP variants.

Designing Reverse Supply Chain Networks with Returned Product Quality Control

Abstract

Ms. Sahar Ebrahimi Bajgani [1], ***Prof. Sara Saberi*** [2], ***Prof. Fuminori Toyasaki*** [3]

1. *Worcester Polytechnic Insti,* ***2.*** *Worcester Polytechnic Institute,* ***3.*** *York University*

For agility and cost-effectiveness, remanufacturing was highlighted as a solution to a capacity shortage for ventilators during the Covid-19 pandemic. Later in post-pandemic times, recycling was introduced as a solution to raw material shortcomings. Adopting these reverse logistics activities can be challenging due to the nature and qualities of reverse flows. There should be a threshold for the quality of returnees to prevent costly remanufacturing. Also, reverse logistics service providers need to obtain the technical information by participating in technology licensing which leads to information leakage. This study's proposal for a reverse supply chain network model includes competing collection centers, 3rd party remanufacturers, and recyclers investigating the case of ventilator products in three pandemic scenarios: Pre-pandemic, During the pandemic, and Post-pandemic. The minimum acceptable quality for remanufacturing is determined. The model is then extended with technology acquisition settings to allow remanufacturers receive OEM licenses for accessing technical needs. The possible leakage of technical information is then considered, along with incentivizing and penalizing strategies. We found that the penalizing strategy improves the technology acquisition level at the cost of reducing the overall profit, while incentivizing strategy can maintain the profit but does not improve the technology acquisition level remarkably.

Determining Production and Distribution Policy in Two-Tier Supply Chain Using Reinforcement Learning

Abstract

Dr. Dmitriy Shaltayev [1], ***Dr. Borga Deniz*** [2], ***Mr. Joona Rahko*** [3]

1. *Christopher Newport University,* ***2.*** *Framingham State University,* ***3.*** *Unity Technologies*

In this paper we use reinforcement learning to determine the optimal production and distribution policy (PDP) for a two-tiered supply chain of a non-perishable item. We model the supply chain as a network consisting of a supplier with random yield and a number of identical retailers. Each retailer observes stationary stochastic demand. The decision making is centralized, and each time period the central decision authority (an "agent") makes two concurrent decisions: how much inventory to ship to each of the retailers, and how many units to make. We assume that both production and transportation lead times are present in the system. Due to the production problems the observed yield from the manufacturing process can be less than the production decision entered into a production queue. If the total of shipment decisions to the retailers exceeds the supplier's on-hand inventory, the missing units are procured from the third-party supplier (spot market) at a higher unit cost, as well as higher shipment cost. Python *ray* library is used to find the optimal PDP.

Drone Routing Problem with Flexible Repeat Visits in Road Traffic Monitoring

Abstract

Dr. Sepideh Alavi [1]

1. *Assistant professor*

This presentation introduces a new version of the vehicle routing problem called drone routing problem with flexible repeat visits that are used in road traffic monitoring. There are some locations on highways that require frequent observations. This problem considers a set of certain locations and a set of available drones. Each location must be repeatedly visited under a maximum inter-visit time requirement, but there is flexibility in the visit times of the locations. The goal is to identify location visit times which satisfy their maximum inter-visit time requirements and create drone routes that minimize the number of drones used. We develop a mixed-integer programming model for this problem and present the results from solving the model by IBM ILOG CPLEX. The CPLEX results for small problem instances indicate that the problem at hand is very complex since CPLEX is not able to find optimal solutions for many instances. As the problem size increases, we see more difficulty for CPLEX in solving the problem.

EXPLORING THE LIMITS OF SUPPLY CHAIN ROBUSTNESS

Abstract

Prof. Henry Adobor [1]

1. *Quinnipiac*

Purpose —This paper aims to explore supply chains as robust-yet-fragile systems. Supply chains may be robust to one set of disruptions, yet remain fragile to others, and the paper seeks to understand how the inherent tradeoffs between robustness and fragility emerge and how we can manage such systems. The robust-yet-fragile framework may allow us to explore the limits of supply chain robustness, resilience, and agility.

Design/approach —Supply chain robustness and fragility literature are used to develop a conceptual framework for exploring the limits of supply chains as designed complex systems.

Findings — One of the hallmark features of robust-yet-fragile systems is that they are good at anticipating threats but terrible at dealing with unanticipated ones. As the complexity of a system grows, both the sources and severity of disruptions increase. Managing the robust-yet-fragile paradox requires the management of the tradeoffs, amongst other things.

Practical implications — Management strategies that aim to reduce uncertainty can sometimes lead to further complexity making such systems more vulnerable to unanticipated disruptions in the long run. Therefore, all strategies for robustness and resilience in supply chains should be viewed as experiments subject to constant adaptation.

Exploring the potential of supply chain business models in the Industry 4.0 era

Abstract

Dr. Pedro Reyes [1], ***Dr. John Visich*** [2], ***Dr. Patrick Jaska*** [3], ***Dr. Michael Gravier*** [2]

1. *Baylor University,* ***2.*** *Bryant University,* ***3.*** *University of Mary Hardin-Baylor*

Industry 4.0 technologies have been rapidly changing the business environment over the past decade. These changes are accelerating how to effectively integrate, deploy, and manage business models in an efficient and effective manner across an organization's supply chain. This paper discusses the supply chain impact of the following technologies: radio frequency identification, the Internet of Things, Industry 4.0, artificial intelligence, machine learning, and blockchains. Based on case study analysis and secondary sources, this paper investigates the business models of selected supply chain models. We identify the implementation challenges, expected benefits of these technologies, and the managerial implications of merging the technologies to create a long-term strategic competitive advantage.

On the Integrated Production and Distribution Problem with Time Windows

Abstract

Dr. Hua Zhong [1]

1. *SUNY Oneonta*

We study a partial LP relaxation based heuristic that schedules the operations of a fleet of heterogeneous vessels in an integrated production and distribution process where the products are released for distribution at different time points. Empirical results and special case analysis on the performance of the proposed heuristic are reported.

Proximity Framework: Supply Chain's Link

Abstract

Dr. David Cortes [1]

1. *Providence College*

Firms that span a wide array of locations, cultures, and industries must interact with one another and make efforts to effectively coordinate their supply chains. Thus, the concept of proximity has become increasingly important in different streams of literature. However, proximity in the context of supply chain represents an important construct that lacks a clear conceptualization and needs further research. In this work, we aim to contribute to the body of literature on proximity as well as that of supply chain by proposing an extension to existing proximity frameworks and incorporating a new dimension of supply chain proximity.

Supply chain resilience and firm innovativeness: an empirical investigation

Abstract

Dr. Sergey Ponomarov [1]

1. *The Citadel*

A dynamic capabilities theory facilitates the discussion on the nature of firm innovativeness and supply chain resilience and informs the development of the proposed conceptual model. Firm innovativeness and supply chain resilience are explored and empirically tested within a supply chain disruptions context. The moderating role of supply uncertainty and interdependence in the focal relationship was also hypothesized and tested. Findings suggest that firm innovativeness is positively associated with supply chain resilience, and supply uncertainty negatively moderates this relationship. This research has both theoretical and managerial implications.
KEYWORDS: Supply chain resilience, supply chain disruptions, innovativeness

The impact of supply chain disruptions and 3D printing investments

Abstract

Dr. Kraiwinee Bunyaratavej [1]

1. *Delaware State University*

When the COVID-19 pandemic hit in the late 2019, the world did not know how it would have the profound impact not only on our lives but also business environment. The new normal emerged including working from home and social distance. In terms of business, the pandemic brought out the vulnerability of the global supply chains since they were greatly disrupted. Shortages of products can be experienced throughout the pandemic ranging from personal protective equipment (PPE) to masks. Some firms opt to increase their capacity. Some business resort to source locally. The 3D printing became an important tool to solve supply chain problems for businesses. In this paper, we will empirically investigated how supply chain problems have an impact on the 3D printing investments.

Marketing and Consumer Behavior

A customer choice approach to retail merchandising: Evidence from the apparel industry

Abstract

Dr. Katherine Ashley [1]

1. *Northeastern University*

In this paper we analyze the optimal merchandising strategy of a fashion retailer with several hundred brick-and-mortar locations. Specifically, the central objective is to identify the profit-maximizing number of unique styles per category for each of the retailer's stores. Using inventory and transaction-level data, we estimate a multinomial logit (MNL) choice model and show that optimal style breadth is increasing in dollar sales volume. Recommendations for unique style count by category at each store are provided and used to create detailed sales plans at the SKU level, which incorporate historical availability of and demand for items in each size. We conclude with insights that can be applied across the retail sector to better allocate inventory across multiple outlets based on the distinct demand patterns of each location.

Does Cancel Culture Affect The Bottom Line? A Semantic Analysis of Cancel Culture Boycotts and Financial Indicators

Abstract

Dr. Paul Reyes-Fournier [1]

1. *Keiser University*

Users generate 500 million tweets per day on Twitter, which translates to 1.85 X 10^{10} words per day. Collectively, these tweets develop ideas with emotional content in the form of sentiment. In the rapidly developing paradigm of cancel culture, it is in the best interest of businesses to understand the effects of this idea propagation on the company's reputation and fiscal position. In this study, the Twitter sentiment of three publicly traded companies that are the target of an online boycott were compared to Altman's *z* score as a predictor of a move toward bankruptcy. Time-series regression analysis was used to develop a predictive model in which the a priori state informs the posterior distribution of correlations to fiscal indicators. Twitter data were collected for each company per fiscal quarter for 10 years, and reduced using a one-dimensional *K* means clustering analysis to give three valenced means (positive, neutral, and negative). Results showed a spike in the sentiment values that corresponded to the call to cancel. Two of the companies showed a resulting structural break in sales, indicating that the boycott was successful. In all three companies, this sudden change in sentiment preceded a shock to the fiscal system of the company, which introduced analytical variance to the system that lasted between two and six periods after the call to cancel. This variance adds error to the predictive value of normal managerial accounting, making this research significant to managers of companies targeted for boycotts.

Estimating Risk of Sex and Labor Enslavement Among Consenting Human Trafficking Job Seekers: a Global Sample

Abstract

Dr. Vernon Murray [1]

1. *Marist College*

Human trafficking entails a variety of situations between victim and trafficker. According to UN and US State Department definitions, some forms of trafficking are by "consent" (i.e., no interpersonal coercion). Chuang (2009), Weitzer (2014), and Murray et. al (2015) in marketing have echoed this. There are also many instances wherein an impoverished person takes a job with a trafficker and is subsequently forced into sex or other work. Countries vary in their percentages of these deceived victims. We sampled eighteen countries and over 25,000 adult victims from the Counter Trafficking Data Collaborative. We computed "Deception Rate" by country. It is offered in this study as an estimate of the relative risk (of enslavement) that an entering, consenting victim makes. Results from our convenience sample indicate that the worst traffickers were in Lebanon (over 90% of initially consenting adults tricked into slavery), followed by Jordan (78%), Russia (62%), the United States (47%), Poland (42%), and Kazakhstan (41%). There was a sharp drop into single digits for United Arab Emirates (9%), Indonesia (6%), and Malaysia (4%). The risk of enslavement also varies by type of human trafficking (sex vs labor) and the victim's relationship with the trafficker. For instance, a person who takes job advice from an "Intimate Partner" is more likely to end up sexually enslaved than someone who takes advice from a "Friend" or "Family Member." Strategic implications are that "Victim Intervention Marketing" (Murray et. al, 2015) interventionists should adapt awareness messages according to consenting victims' risk of subsequent enslavement.

I Don’t Know What That Is: The Effects of Brand Incongruence on Consumer Choices

Abstract

Dr. Rusty Stough [1]

1. *University of Maine*

With brand sponsorships and stadium naming rights being big business (Cornwell et al., 2006), in this paper, we explore the incongruence between stadium naming rights and consumer choices. Through a series of 3 studies, we explore the impact of different types of brand incongruence (Ng et al., 2021) on consumers’ brand perception and choice. Importantly, our findings are consistent with previous research on naming rights research that finds consumers are more likely to remember the brand. However, we find that for geographic incongruence (i.e. a company headquartered outside of the stadium’s state), there is a negative impact on both brand perception and consumer brand choice. In other words even though consumers are more likely to recall these brands, they are more likely to have a negative feeling toward them and be less likely to purchase products from these brands. Further, we find that these feelings can be mitigated if the incongruence can be explained. For example, a company that is headquartered elsewhere, but is opening a new site in the city. We also explore additional types of incongruence such as sport incongruence and demographic incongruence and find similar results. Finally, we explore the impact of incongruence on the sports team itself. Surprisingly, we find little impact on consumer perception of the team. These findings provide important implications for companies who are considering a team sponsorship and provide additional understanding to the role of stadium naming on consumer perception.

Price Perception in Foreign Currencies: The Reference Dependence View

Abstract

Dr. vikas goyal [1]

1. *indian institute of management indore*

Prior research suggests that price perceptions in foreign currencies are primarily guided by the nominal (numerical) value of prices rather than its 'real-value'. Consistent with reference dependence view (Tversky and Kahneman 1991), we demonstrate that individuals use salient reference-price in home currency (i.e. real value) to form price perceptions in unknown foreign currencies. A series of experiments across multiple product categories provide support for our theorization regarding the differential operation of memory-based and stimuli-based reference prices. We find that this effect is moderated by individuals' level of familiarity with the product category. Managerial and theoretical implications of the research are discussed.

Authors Index