Northeast Decision Sciences Institute

Alexandria, Virginia

March 26-28, 2010

Neset Hikmet
Program Chair
Nicholls State University

Albert E. Avery
Proceedings Editor
Towson University

39th Annual Meeting

PLANNING THE OUTSOURCING PROJECT FOR XBRL IMPLEMENTATION

Deb Sledgianowski, Department of Accounting, Taxation, and Legal Studies,
Hofstra University, Hempstead, NY, (516) 463-4759, Deb.Sledgianowski@Hofstra.edu
Nathan Slavin, Department of Accounting, Taxation, and Legal Studies,
Hofstra University, Hempstead, NY, (516) 463-5690, Nathan.Slavin@Hofstra.edu
Robert Fonfeder, Department of Accounting, Taxation, and Legal Studies,
Hofstra University, Hempstead, NY, (516) 463-6988, Robert.Fonfeder@Hofstra.edu

ABSTRACT

Publicly-held companies in the U.S. are now being required by the Securities Exchange Commission (SEC) to file their financial reports using the tagging standard of Extensible Business Reporting Language (XBRL). This interactive data requirement is being phased-in beginning in 2009, with all U.S. publicly-traded companies complying by 2011. The objective of our work-in-progress paper is to discuss eight characteristics of client-vendor relationships organizations should consider when developing a plan for outsourcing implementation of XBRL filing. The eight characteristics are (1) client-vendor relationship, (2) company size, (3) outsourcing contract, (4) intellectual property, (5) communication, (6) custom functionality, (7) task complexity, and (8) vendor's skills.

Keywords: Accounting Information Systems, Project Management

WHAT IS XBRL?

XBRL is a standard open source markup computer programming language that uses header tags to label each individual element of data to enable them to be identified by other software applications. XBRL can be used to electronically communicate business and financial information both intra- and inter-organizationally. XBRL has recently come into the spotlight as an information technology to facilitate the transmission of financial statements that publicly held companies file with the Security Exchange Commission (SEC). For example, when using XBRL to electronically file its 10-K and 10-Q statements, a company would attach a tag labeled "usfr-pte:CommonStockParValuePerShare" to its U.S. GAAP Common Stock Par Value Per Share figure on its interactive data balance sheet submission to signify that the data following the tag is the numerical amount of the Common Stock Par Value Per Share from its financial statement.

IMPLEMENTING THE SEC'S MANDATE

The SEC's interactive data ruling affects over 10,000 domestic and foreign issuers that prepare their financial statements under U.S. GAAP or IFRS. The SEC estimates that preparation costs to a company submitting its interactive data financial statements with detailed footnotes and schedules could average between $21,075 and $30,700. The SEC estimates the number of hours required to tag face financials averages 125 hours for the first submission and 17 hours for subsequent submissions. One of the implementation options that firms are contemplating is to outsource their XBRL filing to a third-party.

Over half of the 50 companies participating in the SEC's voluntary program for XBRL filing reported that they outsourced their XBRL activities [3]. The following eight characteristics (see Figure 1) should be taken into account when considering outsourcing implementation of the interactive data mandate.

CHARACTERSTICS OF OUTSOURCING RELATIONSHIPS FRAMEWORK

1. Client-Vendor Relationship: Firms should consider whether they already have an established relationship with an XBRL service provider. Their accounting firm, financial printer, or provider of financial statement conversion to HTML may also provide XBRL conversion services. It may be more beneficial to continue an established relationship than to develop a new one. Firms should be aware that interactive data will be subject to the same liability under federal securities laws as the corresponding portions of the traditional format filing, and as such, they are still responsible for the content even if they don't actually create the XBRL documents.

2. Company Size: Firms should consider their size relative to the size of the outsourcing vendor. Firms tend to engage services of similar sized firms, especially small-to-medium sized firms. This may be due to difficulties that small firms have in reaching out to larger vendors. Smaller firms may prefer to work with vendors similar in size because they may perceive that they receive better service from them than from larger vendors, who may give preference to more lucrative contracts. When using vendors of similar size, the ability to monitor the vendor's performance and reduce the risk of non-compliance is enhanced [12]. While no empirical research to date has conclusively established a significant advantage from outsourcing client-vendor size parity, the prescriptive literature suggests that small-to-medium sized businesses tend to use vendor companies of a similar size [4] [8]. Some, however, argue that this is because SMBs have difficulties in reaching out to large IT outsourcers [13].

3. Outsourcing Contract: Best practices suggest the use of a formal detailed and comprehensive contract for outsourced services [2]. A level of trust established by a pre-existing client-vendor relationship can reduce the transaction costs inherent in a relationship between two parties. If one side doesn't trust the other side, then resources that could be utilized elsewhere may be consumed to manage the risk that the other party won't keep its side of the agreement. Social network and prior favorable relationships can help reduce the transaction cost of having to prepare highly comprehensive contracts, and "act as substitutes for formal institutional support" [3].

4. Intellectual Property: Transaction-cost theory informs us that, the collaboration between a client and its outsourcing partner creates an asset whose ownership must be agreed upon by both sides [1]. Firms deciding to outsource will still accountable for the content of their interactive data so they may want to consider whether the outsourced vendor offers Statement on Auditing Standards 70 or another form of assurance of their internal controls for safeguarding data and other content.

5. Communication: Research has shown that clear communication of objectives and information exchange can lead to a higher level of team commitment and performance [9].

6. Custom functionality: Unless an organization has unique filing requirements, there probably isn't any custom functionality required to fulfill its XBRL obligations. The common wisdom in outsourcing is to develop a solid business case for *not* implementing a commercially available solution [7] [10]. If one of the existing XBRL software packages can easily be used to solve the business requirements, then it can be a much faster and cheaper solution than developing a solution in-house or outsourcing implementation. If a process is not seen as a core competency, which in many cases financial reporting isn't, then a firm can make a good case for outsourcing its filing.

7. Task Complexity:
Traditional information technology outsourcing involved highly structured application development. The trend now is toward outsourcing less structured projects requiring greater client contact, clarity of outsourcing mission, and project management [6] [14] [11]. Organizations may want to initially outsource XBRL filing if they consider the implementation to initially be complex but as they become more comfortable with the process, they may want to consider other alternatives, for example, an in-house implementation of filing using XBRL software applications.

8. Vendor's Skills: Organizations considering outsourcing should make sure the vendor they use has experience with industry specific XBRL requirements. If special tagging is needed, does the vendor have experience with this? One of the main reasons of outsourcing is to take advantage of the vendor's technical skills [1]. As Gonzalez et al [5] asserted, in the case of IS outsourcing if the provider does not fully understand what the business is all about the customer's needs may not be properly met.

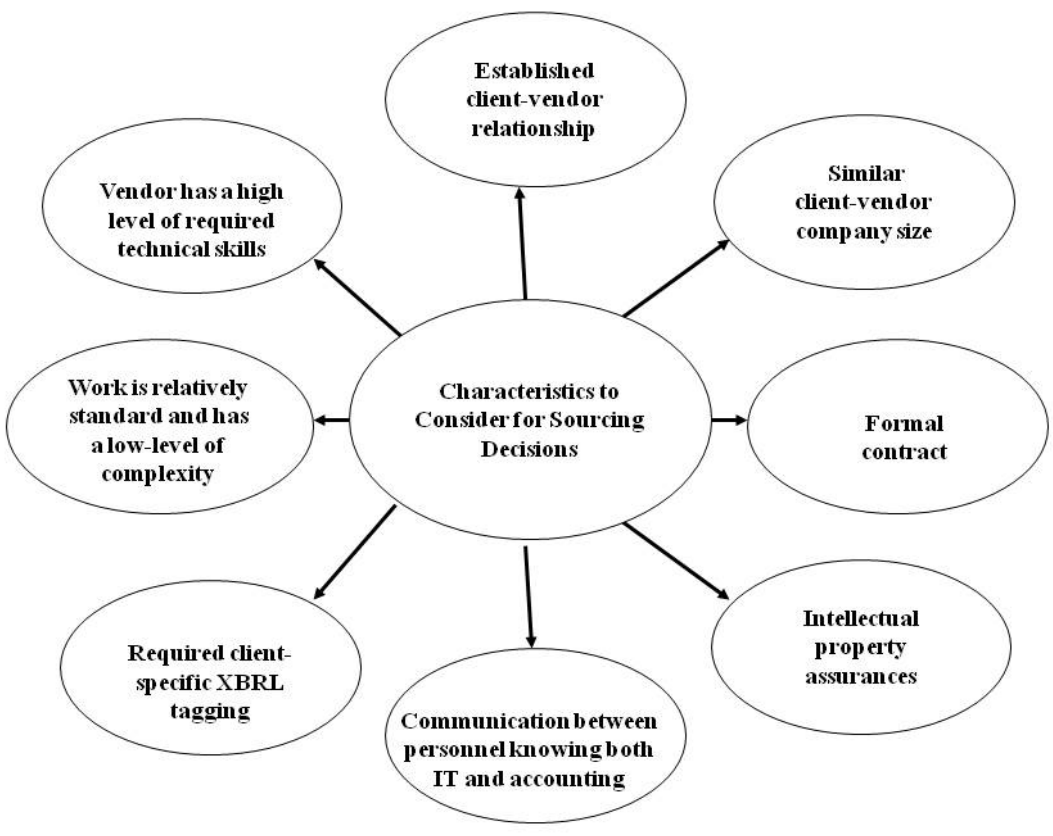


Exhibit 1. Characteristics of XBRL Client-Vendor Outsourcing Relationship

CONCLUSION

Organizations outsourcing their implementation of the recent mandate by the SEC requiring XBRL filing should consider (1) the client-vendor relationship, (2) company size, (3) the outsourcing contract, (4) intellectual property assurances, (5) communication, (6) custom functionality, (7) task complexity, and (8) vendor skills.

ACKNOWLEDGEMENTS

This research was sponsored by a Summer Research Grant from the Frank G. Zarb School of Business at Hofstra University.

REFERENCES

[1] Aubert, B. A., Rivard S., Patry, M. "A Transaction Cost Model of IT Outsourcing", *Information & Management*, 2004, 41(7), pp. 921-932.

[2] Bryson, K.M., Sullivan, W. E. (2003), "Designing effective incentive-oriented contracts for application service provider hosting of ERP systems", *Business Process Management Journal*, 2003, 9(6), pp. 705–721.

[3] Choi, V., Grant. G., Uiz, A. "Insights from the SEC's XBRL Program", *The CPA Journal*, 2008, December.

[4] Fox, G. "What You Should Know about Outsourcing Contracts," *CIO Magazine*, January, 2006.

[5] Gonzalez, R., Gasco, J., Llopis, J. "Information Systems Offshore Outsourcing", *Industrial Management & Data Systems*, 2006, 106(9), pp.1233-1248.

[6] Jennex, M. E. and Adelakun, O. "Success Factors for Offshore Information System Development", *Journal of Information Technology Cases and Applications*, 2003, 5(3), pp. 12-31.

[7] Keil, M. and Tiwana, A. (2006), "Relative Importance of Evaluation Criteria for Enterprise Systems: A Conjoint Study", *Information Systems Journal*, 2006, 16(3), pp 237-262.

[8] Kobayashi-Hillary, M. "It's Time to Focus on Quality and Information Security", October, 2006, available at: www.indiaonestop.com/face2face/mark.htm (accessed on November 11, 2009).

[9] Liang, T., Liu, C., Lin, T., Lin, B. "Effect of Team Diversity on Software Project Performance", Industrial Management & Data Systems, 2007, 107(5), pp. 636-653.

[10] Lucas, H. C. *Information Technology: Strategic Decision Making for Management*, NY: Wiley, 2005.

[11] Mirani, R. "Client-vendor relationships in offshore applications development: an evolutionary framework", *Information Resources Management Journal*, 2006, 19(4), pp.72-86.

[12] Oh, W., Gallivan, M. J., Kim, J. W. “The Market's Perception of the Transactional Risks of Information Technology Outsourcing Announcements”, *Journal of Management Information Systems*, 2006, 22(4), pp 271-303

[13] Paul, A. (2007). “Size Impedes SMEs' Tryst with Outsourcing”, *The Economic Times*, March 16, 2007.

[14] Xue,Y., Sankar, C. S., Mbarika, V. W. “Information Technology Outsourcing and Virtual Teams”, *Journal of Computer Information Systems*, 2004/2005, 45(2), pp. 9-16.

Learning Through Role Playing

Fraud is becoming a common practice in the Internet age putting pressure on universities to include fraud investigation as part of their business curriculum. This paper describes role playing as a learning tool in a accounting forensics course . According to Wikipedia, "*forensics* encompasses the accepted scholarly or scientific methodology and norms under which the facts regarding an event, or an artifact, or some other physical item (such as a corpse, or cadaver, for example) are to the broader notion of authentication whereby an interest outside of a legal form exists in determining whether an object is in fact what it purports to be, or is alleged as being." The inter-disciplinary nature of the course makes it quite challenging, but an interesting topic to teach. The next section describes the experiment.

The Experiment

The study involves an undergraduate class that is required of all accounting majors at a mid-eastern university. A typical class has 30 to 50 students depending on the delivery mode. Class size is limited to 30 students in an on-line class, but it is allowed to go up to 50 in a face-to-face class. A problem related to missing accounting data was developed. Students were provided with adequate information related to digital and accounting forensics through lectures and videos. Students were allowed to create their own group and select role-playing. Groups of four students were encouraged so that each student could cover one part of the accounting forensic investigation. Students were advised on the concept of role playing and how it works. Students were asked to watch a show and were asked to simulate it in the class and then relate it to exercise at hand.

Analysis and Results

The experiment revealed that almost thirty three percent of students were able to map role playing to forensics. is continuing this semester. Results were measured by their performance in role-playing and the final case analysis.

Conclusion

Accounting forensics is becoming important as the economy moves into recession and more and more computer-based fraud is committed. Many functional areas have come under scrutiny for fraud, embezzlement of funds, information gaps, etc., and many are committed to using information technologies. As the Internet continues to grow, so does Internet based fraud. We must teach our students on ways to recognize and prevent fraud. This paper is an attempt in that direction.

References

provided on request

DOES PERFORMANCE IN A STUDENT'S FIRST ACCOUNTING COURSE PROVIDE INSIGHT INTO PERFORMANCE IN OTHER BUSINESS CORE COURSES?

Richard L. Baker, Bloomsburg University, 570-389-4561, rbaker@bloomu.edu
William E. Bealing, Jr., Bloomsburg University, 570-389-4386, wbealing@bloomu.edu
Edward Pitingolo, Bloomsburg University, 570-389-4122, epitingo@bloomu.edu

ABSTRACT

This study examines the relationships between student performance in the initial financial accounting course taken by business majors and their performance, as measured by course grades, in the other non-accounting business core courses. For accounting majors, performance in the first financial accounting course was found to have a statistically significant relationship to performance in all of the non-accounting core courses except Business Policies. For the other business majors, there was a significant relationship between performance in the first accounting course and all of the business core courses except the introductory marketing course, the core computer course, and Business Policies.

The results indicate that a student's performance in their first financial accounting course is a valid predictor of his/her success in most non-accounting business core courses. This raises the possibility that the content of the introductory accounting course could require students to think critically and, therefore, a student's performance in this course could be a proxy for critical thinking abilities brought to the other business core courses.

INTRODUCTION

What factors determine an individual's level of success in completing his/her business core courses? A student's ability and motivation are often cited as key determinants. However, these traits have proven hard to measure in quantifiable terms. Prior research has examined the link between personality type and accounting performance [3] [10] [11] [12]. It was found that a specific personality type, sensing-judging (SJ) is related to successful performance, as measured by course grade [3]. The authors hypothesized that personality type could be used as a surrogate measure of such unquantifiable characteristics as motivation and ability. However, could there be other readily available indicators that would accurately predict a student's performance in a course?

This paper extends these earlier lines of research by examining the theory that the grade received in one accounting course can predict performance in subsequent non-accounting courses. Once a student completes a course, their grade becomes new information which can be used to predict future performance. In the spirit of Bayesian statistics, this new information is used to "update" the assessment of a student's performance in a subsequent accounting course. Researchers [4] have found that grades earned in prerequisite accounting courses were statistically significant predictors of performance in subsequent accounting courses. This paper extends this line of research by positing that a student's performance in their introductory accounting course should be positively correlated to performance in subsequent non-accounting business core courses.

This logic is intuitively appealing because, like the introductory accounting course, most business core courses purport to require students to be able to think critically, rather than to perform rote memorization. If this is true, a student's performance in the introductory accounting course should demonstrate a positive correlation to his/her ability to succeed in other business core courses.

Other researchers have also found that student performance in certain courses was associated with performance in specific later courses. Various authors [1] [5] found that students who had taken calculus and related mathematics courses performed better in economics and finance courses than individuals who had no prior coursework in the field of mathematics. Others [6] concluded that completed accounting courses were significantly correlated with student performance in other business courses.

In the area of accounting, prior research found that students' grade point averages (GPA) and performance in their accounting principles course, as well as performance in their first statistics course, were all significantly correlated to their success in cost accounting [8].

It is interesting to note that one group of authors found that prior related coursework does not predict performance in college level information systems (IS) courses [9]. Their research looked at students who had IS courses in high school and correlated it to later performance in an introductory college level IS course. Their results are consistent with one of the first studies to specifically look at achievement levels in high school accounting courses and subsequent performance in the first college level accounting course. Another group of researchers found no difference in the performance of students who had bookkeeping courses in high school versus those who did not [2]. However, another set of researchers found pre-college study of accounting/bookkeeping was significantly related to subsequent performance in an introductory accounting course [7].

METHODOLOGY

This study is based on the performance of students in business core courses at one, mid-sized, public university. The research began with a review of the list of non-accounting business core courses required at that institution. Ten courses were identified and used as the basis for testing of:

> **H1** There is no significant positive correlation between performance, as measured by course grade, in the first financial accounting course and performance, as measured by course grade, in subsequent non-accounting, business core courses.

The ten courses which were identified were: 1) Business Law I, 2) Principles of Economics I, 3) Principles of Economics II, 4) Business & Economics Math, 5) Business and Economics Statistics, 6) Matrix Algebra, 7) Principles of Marketing, 8) Principles of Management, 9) Information Technology, and 10) Business Policies.

The letter grades received by students on their first attempt at each of the courses, which are required of all business majors, were recoded into numerical values. The following scale was used: A=4; B=3; C=2; D=1; E=0. Grades of W (withdrew), I (incomplete), and TR (grade transferred in) were not considered in the initial analysis. Subsequently, all of the analyses were repeated with grades of W included and assigned a value of 0. However, no statistical differences in the results were observed.

Although the university employs a plus/minus grading system the determination was made to disregard any plus or minus in the grade. For example, if a grade of C+ was earned, it was treated as a "C" for purposes of this study. This was done because not all faculty members award plus/minus grades even though university policy permits the practice.

Regression models were employed to examine the relationship between individual student performances in the first financial accounting course and subsequent non-accounting, business core courses.

In an effort to expose any bias based upon the students' natural intelligence, the entire analysis process was subsequently repeated substituting self-reported student total SAT scores for the grade received in the first financial accounting course. The same ten business core courses were used for testing of:

H2 There is no significant positive correlation between total SAT score and performance, as measured by course grade, in subsequent non-accounting, business core courses.

RESULTS

Table 1 contains a summary of the significance (p-values) of the regression analysis of the performance of accounting majors in the ten non-accounting business core courses. The p-values for simple regression models using performance in the first financial accounting course, SAT scores, as well as for multiple regression models using both variables, are reported for all of the core courses. All of the regressions involving the introductory financial accounting course yielded significant results in the anticipated direction except for the Business Policies course. The R-Squared values ranged from a low of .0743 for Business and Economics Math to a high of .2934 for Matrix Algebra.

At this point, the models were re-specified to examine the relationship between a student's SAT score and his/her performance in the ten business core courses under examination. This was done to examine the validity of the argument that "better students do better in their classes." A student's total SAT score was used as a proxy for the student's innate abilities.

The SAT scores proved to be significant only for the Law, Economics I, Economics II, and Matrix Algebra courses. As the table also indicates, during multiple regression, only Economics I and Matrix Algebra remained significant for <u>both</u> variables.

TABLE 1
Accounting Majors

Non-Accounting Course	Grade=fcn (Fin) P-Value	Fin R-Squared	Grade=fcn (SAT Score) P-Value	Grade=fcn (Fin and SAT)
Law	<.0001	.1967	.0299	Prin I Signif
Economics I	<.0001	.1524	<.0001	Both Signif
Economics II	<.0001	.2515	.0126	Prin I Signif
Business and Economics Math	.0008	.0743	Not Significant	N/A
Business and Economics Statistics	.0016	.0874	Not Significant	N/A
Matrix Algebra	<.0001	.2934	<.0001	Both Signif
Principles of Marketing	.0004	.1154	Not Significant	N/A
Principles of Management	<.0001	.1421	Not Significant	N/A
Information Technology	<.0001	.1515	Not Significant	N/A
Business Policies	Not Significant	N/A	Not Significant	N/A

TABLE 2
Non-Accounting Majors

Non-Accounting Course	Grade=fcn (Fin) P-Value	Fin R-Squared	Grade=fcn (SAT Score) P-Value	Grade=fcn (Fin and SAT)
Law	<.0003	.1226	.0299	Fin Signif
Economics I	<.0001	.1560	<.0001	Fin Signif
Economics II	<.0001	.1597	.0126	Fin Signif
Business and Economics Math	<.0001	.2791	Not Significant	N/A
Business and Economics Statistics	.0201	.1542	Not Significant	N/A
Matrix Algebra	.0092	.0748	<.0001	Fin Signif
Principles of Marketing	Not Significant	N/A	Not Significant	N/A
Principles of Management	.0012	.2847	Not Significant	N/A
Information Technology	Not Significant	N/A	Not Significant	N/A
Business Policies	Not Significant	N/A	Not Significant	N/A

Table 2 contains a summary of the significance (p-values) of the regression analysis of the performance of non-accounting majors in the ten non-accounting business core courses. The p-values for simple regression models using performance in the first accounting course, SAT scores, as well as for multiple regression models using both variables, are reported for all of the business core courses. All of the regressions involving the introductory accounting course yielded significant results in the anticipated direction except for the Business Policies course. The R-Squared values ranged from a low of .0748 for Matrix Algebra to a high of .2848 for Principles of Management.

As the table indicates, there was a significant positive relationship between performance in the first accounting course (Financial) and all of the courses except Principles of Marketing, Information Technology, and Business Policies. However, only the Law, Economics, and Matrix Algebra courses proved to have a significant relationship with student SAT scores. Under multiple regression, none of the courses proved to be significant for both variables.

DISCUSSION AND IMPLICATIONS

The null hypothesis **H1** was rejected for all business core courses taken by accounting majors except Business Policies. For all non-accounting majors the null hypothesis was rejected for all core courses except Principles of Marketing, Information Technology, and Business Policies. Thus, the results of this research support the contention that performance in a student's first accounting course provides insights into their performance in other non-accounting business core courses. Students who do well in their initial accounting course also tend to do well in later non-accounting business core courses. Likewise, a student who does not perform well in the first accounting course tends to have difficulty in subsequent business core courses.

Interestingly, the null hypothesis **H2** can only be rejected for four courses, Law, Economics I and II, and Matrix Algebra. For the other six courses there was no significant correlation. These relationships applied to both accounting and non-accounting majors.

A few possible explanations emerge. First, it could be that SAT scores are proxies of the students' innate abilities, but say little about their willingness to exert effort. It could also be that students' grades in their financial accounting course provide additional information about their willingness to work hard. Many times students who work hard, but possess lower levels of innate ability, outperform their "brighter" counterparts.

Second, the possibility exists that material contained in the introductory financial accounting course requires higher level thought, above and beyond memorization, and provides a proxy for a student's critical thinking skills. For years, accounting educators have said that their courses required higher level thought, not merely memorization of specific entries and statement formats. The results of this study appear to provide some support for that position.

SUMMARY AND FUTURE RESEARCH

In all of the instances examined in this study, a significant relationship was found between performance in a student's first financial accounting course and subsequent non-accounting business core courses except, Business Policies. Future research needs to be conducted to determine if the course content of the introductory accounting course could be used as a construct for critical thinking abilities. It could be that the initial financial accounting course helps to develop important critical thinking skills which students then employ in other business courses. In addition, further research needs to be conducted at other universities to determine if the results of this study are institution specific, or can be generalized.

REFERENCES

[1] Anderson, G., and D. Benjamin, (1994). The Determinants of Success in University Introductory Economics Courses, *Journal of Economic Education*, 25, 99-118.

[2] Baldwin, B.A. and K.R. Howe, (1982). Secondary-Level Study of Accounting and Subsequent Performance in the First College Course, *The Accounting Review*, 57 (3), 619-626.

[3] Bealing, W.E., A.B. Staley and R.L. Baker, (2008). Success Breeds Success: Why Course Prerequisites Matter, Unpublished working paper presented at NEDSI.

[4] Bealing, W.E., R.L. Baker, and C.J. Russo, (2006). Personality: What It Takes To Be An Accountant, *The Accounting Educators' Journal*, 16, 119-128.

[5] Didia, D. and B. Hasnat, (1998). The Determinants of Performance in the University Introductory Finance Course, *Financial Practice and Education*, 8 (1), 102-107.

[6] Ely, D.P. and L. Hittle, (1990). The Impact of Performance in Managerial Economics and Basic Finance Course, *Journal of Financial Education*, 8 (2), 59-61.

[7] Eskew R.K. and R.H. Faley, (1988). Some Determinants of Student Performance in the First College-Level Financial Accounting Course, *The Accounting Review*, 63 (1), 137-147.

[8] Kirk F.R. and C.A. Spector, (2006). Factors Affecting Student Achievement in Cost Accounting, *Academy of Educational Leadership Journal*, 10 (1), 91-104.

[9] Kruck S.E. and D. Lending, (2003). Predicting Academic Performance in an Introductory College-Level IS Course, *Information Technology, Learning, and Performance Journal*, 21 (2), 9-15.

[10] Nourayi, M.M., and A.C. Cherry, (1993). Accounting Students: Performance and Personality Types, *Journal of Education for Business*, November/December, 111-115.

[11] Oswick, C. and P. Barber, (1998). Personality Type and Performance in an Introductory Level Accounting Course: A Research Note, *Accounting Education*, 7 (3), 249-254.

[12] Ott, R.L., M.H. Mann, and C.T. Moores, (1990). An Empirical Investigation into the Interactive Effects of Students Personality Traits and Method of Instruction (Lecture or CAI) on Student Performance in Elementary Accounting, *Journal of Accounting Education*, 8, 17-35.

THE IMPACT OF CONSUMER MOOD ON BRAND EXTENSION EVALUATION

Arthur Cheng-hsui Chen, National Yunlin University of Science and Technology, Taiwan, chencs@yuntech.edu.tw
Shiou-huei Chiou, National Yunlin University of Science and Technology, Taiwan, g9422736@yuntech.edu.tw

ABSTRACT

The objective of this study is to examine the positive/negative mood effects on moderate and near brand extension respectively, moderating by brand concept, gender and age. Two studies were conducted using between subject design experiments. Undergraduate students and the fifth grade elementary student were recruited as research subject via convenient sampling. The evidence supports that utilitarian brand will be enhanced by the positive mood in the moderate distance extension, but not for hedonic brand. For negative mood, its negative impact will also happen for utilitarian brand in the close distance extension, but not for hedonic brand. The results also found that there were moderating effects for age and gender for both positive and negative mood.

Key words: brand extension, mood, category similarity, brand concept, gender, age

INTRODUCTION

The significance of brand extension has been strengthened by the academicians and practitioners in the past ten years. Prior research of brand extension has mostly adopted categorization theory to evaluate the possibility of success. According to logic reasoning and empirical studies, the category similarity or the fit between core product and extended product has been recognized as the key successful factor for brand extension (Volckner and Sattler, 2006). However, there are some successful examples with far extension, such as YAMAHA pianos (core product) and YAMAHA motorcycles (extended product).

Recently, mood has been introduced into this field of brand extension to examine its effect on the similarity perception of brand extension (Barone & Miniard 2002; Barone et al., 2000; Barone 2005). Barone et al. (2000) observe that positive mood would decrease customers' cognitive capacity. Therefore, customers would evaluate the product extension category with mild attitude when they have positive mood. Then, they will enhance the similarity between core product and extended product and create the transfer of brand evaluation. Positive mood enhances evaluations of extension only occurs to moderate brand extension, not in far extension. Barone and Miniard (2002) probe deeply into the preference brands of customers. The result shows it exists with favorable brands except for unfavorable brands. Barone (2005) uses customer involvement to explore the moderating effect. Mood has a positive effect in high but not in low involvement conditions. Yeung and Wyer (2005) focus on product attributes and show that customer mood would be a significant factor to influence the evaluation of extension when expose to favorable brands.

Drawing on the work of Barone and his colleagues, the results of investigating the impact of positive mood on brand extension are robust. Based on the literature, most of the researches regarding the brand extension are centered on the positive mood. If applying negative mood into the research, would there be different for the evaluation of brand extension between positive and negative mood? Of particular interest to this study is to examine the factors moderating the relationship between mood and brand extension evaluations. For example, the different brand concepts have different extendabilities and performance (Park, Milberg, and Lawson 1991). Therefore, the first moderator- utilitarian-oriented or hedonic-oriented brands would be taken into consideration. In consumer research, different gender will have different purchase behavior (Pahl 1990；Wood 1998). Meyers-Levy and Sternthal (1991) found gender differences in levels of cognitive elaboration, which may influence on consumer's information processing. The center argument of mood impacting on extension evaluation is resulted from information processing. Hence, this study would like to examine the gender's moderating effect on the relationship between mood and extension evaluation. Zhang and Sood (2003) found that there are different evaluations of brand extension between adults and children. If consumer judgment or perception on extension will be different between adult and children, then we expect this might happen in the different mood effects. So, the third moderating factor is consumer age.

In summary, this study will use these three moderating factors to examine the enhancement effects of positive mood on moderate brand extension, which normally be evaluated lower than near extension. On the other hand, this study will also focus on the negative impact of the negative mood in the near extension, which has been confirmed as the most successful situation.

LITERATURE AND HYPOTHESES

Brand extension

A general finding from the brand extension literature is that evaluative transfer occurs is related positively to the perceived similarity between the parent brand and extension (e.g., Aaker & Keller, 1990; Boush & Loken, 1991; Boush et al., 1987). From theoretical perspective, categorization theory has frequently been invoked in modeling the brand extension evaluation process. Specifically, consumers make an initial determination by assessing the extent to which the extension belongs to the parent brand category. When such categorization is successful, they simply transfer their evaluation of the parent brand to the extension. However, when the extension is not easily assigned to the category membership, consumers would try other available information like product attribute to perform its judgment. Thus, consumers would be quite likely to transfer their preference for the parent brand to "near" extensions involving products viewed as highly similar to those currently marketed by the brand. Conversely, such evaluative transfer would become more difficult if the extensions distance were considered moderate or much longer to the parent brand. Based on preceding discussion, we can conclude that any factor that improves the perceived similarity of the parent brand and the extension will increase the likelihood that parent brand attitude will be transferred to the extension.

Mood and information processing

Mood is normally defined as a consumer's affective state that is relatively global in nature, as opposed to emotions, which tend to have a specific cause (Gardner, 1985; Luomala & Laaksonen, 2000; Rusting, 1998). Some studies have explored how moods impact on consumer behavior like recall (Lee & Sternthal, 1999), shopping intentions (Swinyard, 1993), the amount of cognitive elaboration engaged in by consumers (Batra & Stayman, 1990), and advertisements (Goldberg & Gorn, 1987; LaTour and LaTour 2009; Martin 2003), and music (Holbrook& Gardner, 2000). Within the field of mood research, a variety of moods have been used like positive moods and happy moods, or negative moods, anxious moods, and angry moods. However, the present study would like to examine the positive and negative moods.

In terms of information processing, research suggest that people engage in heuristic processing during positive moods, and more effortful processing during negative moods. Two perspectives, cognitive capacity and mood-as-information, can be used to explain this phenomenon. Specifically, people with positive moods will have less cognitive capacity available for processing, which results in a greater likelihood of heuristic processing. On the other hand, positive moods also provide a happy/pleasant environment, where indicates that there is little need for elaborative processing. By contrast, negative moods usually come with a more problematic environment, which may facilitate people to be more thinking and elaborative processing.

Regarding consumer judgment, most of the research posits that mood states influence judgments in a mood-congruent manner. This indicates that mood states prime the recall of memories of a similar affective valence (Bower, 1981). Specifically, positive moods prime positive memories; negative moods, negative memories. Accordingly, positive moods result in more favorable evaluations whereas negative moods result in more negative evaluations (Luomala & Laaksonen, 2000). However, some studies have also found mood-incongruent results where negative moods result in favorable evaluations (e.g., Erber & Erber, 1994; Rusting & DeHart, 2000). Furthermore, some research suggests that the role of negative moods is unclear in terms of when and why mood-congruent or mood-incongruent effects will occur (e.g., Rusting, 1998; Rusting & DeHart, 2000). Thus, studying these mood states offers scope to contribute to the understanding of mood regulation, regarding the improvement of a sad mood to a happier mood state.

Mood and brand extension

Recently, mood has been introduced into this field of brand extension to examine its effect on the similarity perception of brand extension (Barone & Miniard 2002; Barone et al., 2000; Barone 2005). A general logic from literature is that positive mood promotes flexibility in categorization and enhances the extent to which a category and an exemplar are perceived as being related to one another. For example, Isen and Daubmun (1984) argue that the influence of mood on the evaluations of extensions introduced by favorably evaluated core brands will depend on its effect on perceptions of the similarity existing between the core brand and the extension. Following from Isen and Daubmun’s (1984) work, Barone, Miniard & Romeo (2000) found that positive mood can enhance evaluation through the perceptions of similarity between the core brand and the extension to a greater extent only for moderate extension but not for near or far extensions. Barone and Miniard (2002) found that there is an asymmetrical effect for mood to the extensions offered by desirable versus undesirable brands. Specifically, positive mood can enhance extension evaluations for moderate extensions introduced by desirable core brands, but not by undesirable core brands. Barone (2005) introduced another factor-involvement to further investigate the impact of mood on brand extension. The results indicate that participants’ mood influenced their evaluations of extensions that were moderately similar to the core brand under conditions of high involvement, but did not happen under low involvement.

H1：The positive mood will have positive enhancement on the evaluation of moderate distance brand extension for utilitarian brand, but not for hedonic brand.

H2：The negative mood will have negative influence on the evaluation of close distance brand extension for utilitarian brand, but not for hedonic brand.

H3：For female, the positive mood will have positive enhancement on the evaluation of moderate distance brand extension, but for male, its effect will not be significant.

H4：For male, the negative mood will have negative influence on the evaluation of close distance brand extension, but for female, its effect will not be significant.

H5：For adult, the positive mood will have positive enhancement on the evaluation of moderate distance brand extension, but for children, its effect is not significant.

H6：For adult, the negative mood will have negative influence on the evaluation of close distance brand extension, but for children, its effect is not significant.

METHODOLOGY

Measurement and Definition

Mood: Mood was the independent variable in this research. Positive and negative mood were manipulated via watching an amusing films and reading the reports of solitude old man and poor children in Africa. According to the Emotion Rating Scale of Izard & Buechler(1997), Yomkins(1980), Watson, Clark & Tellegen (1988), four measurement items in 7 points semantic scale including "Extremely bad mood to extremely good mood", "Sad to happy", "Depressed to Delighted" and "Sorrowful to Cheerful" were used to check the manipulation of mood.

Similarity: Similarity means cognition distance of brands' core extension and extended extension to customers. In terms of cognition distance, we adopted the classification method of Barone, Miniard & Romeo (2000) and it was divided into near and moderate extension. There were two kinds of questionnaires, utilitarian product and hedonic product. Designing the advertisements of products' near and middle extension is the next step. The main purpose was to choose products in near or moderate extension condition. The similarity survey had two parts, utilitarian products and hedonic products.

Perceived quality: Perceived quality represents customers' cognition level of total quality to a product (Dodds, Monroe, and grewal, 1991; Grewal, Monroe, and Krishnan, 1998; Zeithaml, 1988). We used "expected quality" to measure perceived quality and Likert 7 point scale were adopted in this part. According to Keller & Aaker(1992), questions conclude "Extremely bad quality to Extremely good quality".

Brand belief: In this research, questions of brand belief were used to measure customers' opinion or thoughts with extended product. Similarly, semantic 7 point scale was used to develop different questions according to different brand concepts (utilitarian and hedonic). Questions of utilitarian products included "extremely functionless to extremely functional", "extremely useless to extremely useful" and "extremely unattractive to extremely attractive". Questions of hedonic products include "extremely unfashionable to extremely fashionable" and "extremely uncreative to extremely creative" and "extremely unattractive to extremely attractive".

***Brand attitude**:* attitude means the demand level of a product that customers expect to satisfy them. The brands attitude of customers could be evaluated by their own and based on products or brands' cognition. Semantic 7 point scale was used in this research and questions include "extremely bad to extremely good", "extremely unfavorable to extremely favorable" and "extremely undesirable to extremely desirable".

Research Design

Experimental design: Two experimental studies were conducted in this study. In study 1, four major variables- mood, brand concept, extension distance, and sex were put into the experimental design. Extensions were separated into near and moderate extensions. Positive mood was aimed at the moderate brand extension and negative mood was aimed at the near brand extension. Neutral mood was the control group. Brand concepts were divided into utilitarian and hedonic. Then, a similar 2 (brand concept: utilitarian/hedonic) x 2 (extension distance: near/moderate) x 2 (mood: positive or negative/ neutral) between-subject experimental design was created. The sex variable was a within-subject design by increasing the amounts of sample to have enough male and female participants in each cell of the experiment. The main purpose was to test hypothesis 1 to 4. Next, the age variable was added in the study 2 via replacing the sample in each experimental cell from university student in study 1 to junior high school student. The results were used to test hypothesis 5 to 6.

Choosing products and brands: Because gender, age and product attributes were the moderate variables, we had to use the well-known and familiar products to female and male (gender difference), graduate students and junior high school students (age difference). Therefore, Motorola mobile phones and Swatch watch were chosen via a pretest to be the representation of utilitarian and hedonic brands respectively. In another pretest, we developed six kinds of advertisements. Likert 5 point scale was used to investigate what participants feel about the similarity level of mobile phones to other extended products include computer, record pen player, watch, MP3, camera and electronic dictionary. Similarly, clock, bracelet, key chain, mobile phone, wallet and sun glass were used to compare the similarity level to watch. Participants had to choose the product categories of near and moderate extension. Then, Moto digital camera (near extension) and Moto electronic dictionary (moderate extension) were created. Swatch clock (near extension) and Swatch bracelet (moderate extension) were created.

Design of questionnaires*:* Contents of questionnaires had three categories, positive, moderate and negative mood questionnaires (experimental group). The first part was the statement of mood manipulation following the manipulation check measurements. The second part was to read the advertisement of brand extension and the third part was the extend product evaluations. Advertisements had near and moderate extensions to match with utilitarian and hedonic products.

Participants: We chose university and graduate student to represent the elder age participants and junior high school students to represent the younger age participants. Total 730 valid samples via convenient sampling were recruited as research subjects.

RESULTS

To investigate the effect of different moods to brand extension, participants' mood has to be manipulated first. We process brand extension evaluations in the condition of different moods (positive, moderate and negative). Results of manipulation check indicate that there are significant difference between positive, moderate and negative moods (P=0.000). Positive mood (M=5.869) has main significant effects (F=1214.9, P=.000) than neutral mood (M=4.827) and negative mood (M=2.317). Cronbach's α was processed to analyze the consistency of all variable measurement items that include mood, brand belief and brand attitude. Cronbach's α value is between 0.8509~0.9756 and has high Reliability.

Accordingly, mood has significant effect to brand extension evaluations (P=0.000) and product extended distance (P=0.044). To examine the cross-interaction, 2 (mood) x 2 (product extended distance) x2 (product attribute) mixed design were conducted. Mood, product extended distance and product attributes have significant effect (P=0.000). Similarly, mood, product extended distance and gender have significant effect (P=0.007). Mood, product extended distance and age have significant effect (P=0.021).

Evaluations in condition of positive mood are better than neutral moods. The results show the significant difference of utilitarian aspect (P=0.004). Although product quality (P=0.082) is not significant, mean of positive mood is higher than neutral mood (4.961>4.701). The results indicate positive mood could enhance the evaluations of utilitarian brand. However, in terms of hedonic brand, there is no significance (P=0.112). Therefore, the positive mood will have positive enhancement on the evaluation of moderate distance brand extension for utilitarian brand, but not for hedonic brand. Then, H1 is supported.

Regarding negative mood effect, the result shows that there had significant effect to utilitarian brand with near extensions (P=0.039). Specifically, perceived quality (4.26 vs 4.73, p=0.046), brand belief (4.41 vs 4.94, P=0.011) and brand attitude (4.28 vs 4.70, P=0.049) all have significant difference between negative and neutral mood. In contrast, as expected, there is no significant difference for the evaluation of hedonic products (P=0.229). This means that, for near extension, negative mood does not have negative effect for hedonic brand, but does have for utilitarian brand. Hence, H2 is supported.

For the moderating effect of sex, the positive mood, for male, had not significant effect on the moderate extension. However, the positive mood (vs. neutral mood) does have significant positive effect (p=0.032) on the extension evaluation. Specifically, the significant effect happened across all dependent variables-perceived quality (5.01 vs. 4.66, p=0.02), brand belief (5.04 vs. 4.56, P=0.007) and brand attitude (4.84 vs. 4.37, P=0.036). This result is consistent with what we have expected. Therefore, H3 is supported.

For the negative effect of negative mood on near brand extension, the different sex had different results. As we expected, the negative mood for male had higher negative effects than that of female. Specifically, for female, there are no significant difference between negative mood and neutral mood. But the significant negative effect (p=0.026) does happen for male. The evaluations of near extension on all dependent variables- perceived quality (4.52 vs. 4.92, p=0.06), brand belief (4.35 vs. 5.04, P=0.002) and brand attitude (4.22 vs. 4.75, P=0.014) for negative mood were significant lower than that of neutral mood. Therefore, H4 is supported.

At last, we examined the moderating effect of age. The results indicate that for adults group, the positive mood had positive effect on the moderate brand extension evaluation across all variables like perceived quality (4.78 vs. 4.43, p=0.020), brand belief (4.60 vs. 4.23, P=0.011) and brand attitude (4.25 vs. 3.93, P=0.031). However, the positive effect of positive mood did not be found for children. Hence, H5 is supported.

Despite the negative effect of negative mood on near extension for male was significant, based on above analysis, but not female, but combining the female and male as adult group, the negative effect was still significant for all items- perceived quality (4.56 vs. 4.90, p=0.018), brand belief (4.40 vs. 4.86, P=0.015) and brand attitude (4.29 vs. 4.62, P=0.029). However, the negative effect was also slightly significant for children. This is basically still consistent with our expectation. Then H6 is supported.

THE STUDY OF EXPLORATORY AND EXPLOITATIVE INNOVATION IN TAIWAN IT INDUSTRY

Yi-Pei Li, National Dong Hwa University,Taiwan, pn1880@ms23.hinet.net
Dahan Institute of Technology,Taiwan,
Yuh-Yuan Tsai, National Dong Hwa University,Taiwan, yytsai@mail.ndhu.edu.tw

ABSTRACT

The purpose of this study is to understand how firms in Taiwan's IT industry build the competencies in OEM business and OBM (own brand management) through exploratory innovation and exploitative innovation. We also examine the impact of some antecedents such as isolation mechanisms, switching cost and asset-specific investment on both types of innovation activities. Statistical results revealed that exploratory innovation had a positive effect on the competencies needed in own brand management, and explorative innovation had a positive effect on the competencies needed in OEM business.
Keywords: exploratory innovation, exploitative innovation

INTRODUCTION

When we review the development process of Taiwan's information technology industry, especially the IT hardware industry, it is amazing that companies in IT industry achieve an impressive imagine in OEM (Original equipment manufacturing) business. But as competition intensified and the pace of change accelerated, these firms need to renew themselves by exploiting existing competencies and exploring new ones (Floyd & Lane, 2000). It means that companies in IT sector should not only improve their manufacturing ability to maintain the competitive advantage of OEM business, but build new ability to create new business model for their long-term viability.

Because of the threat by the lower labor cost of China in OEM, many companies start to develop their own brand. They adopt OBM (Own brand management) strategy to look for new market opportunities. However, OEM strategy and OBM strategy are different business models and they need different competences to create or maintain their competitive advantage. Firms that engaged in OEM business have to improve their ability in manufacturing efficiency and product quality. Firms that engaged in OBM have to improve their ability in marketing and product design. To achieve these abilities, firms need to engage in different innovation activities.

The notion of exploration and exploitation has emerged as an underlying theme in research on organizational learning and strategy (Levinthal & March, 1993), innovation (Lee et al., 2003; Rothaermel & Deeds, 2004), and entrepreneurship (Shane & Venkataraman, 2000). In this study, we treat exploration and exploitation as two types of activities of innovation. The purpose of this study is to understand how companies in IT industry build the competences needed in OEM and OBM business through exploratory and exploitative innovation. Research on exploratory and exploitative innovation also shows that the antecedents of both activities are different. Therefore, we also want to find out what characteristic in Taiwan's IT industry will influence the type of innovation activities they pursuit.

But the profit earned by OEM business is smaller today, due to the intensive competition from China. We suggest that OEM Manufacturers in Taiwan need to find another business model for long-term survival. Therefore, entering new markets or building own brand may be a good alternative for these companies. When OEM manufacturers pursue exploratory innovation to enter new market or build their own brand, they must ensure that the competitive advantage and profit they own today will not imitate easily by their competitors. We believe that a set of strong isolating mechanism will solve this problem.

HYPOTHESES

March (1991) introduced exploration and exploitation as follows. Exploration includes things captured by terms such as search, variation, risk taking, experimentation, flexibility, discovery and innovation. Exploitation includes such things as refinement, choice, production, efficiency, selection, implementation, and execution (March, 1991). Therefore, exploration is variation-seeking, risk-taking and experimentation oriented. Exploitation is variety-reducing and efficiency oriented. These two concepts require different structure, process, strategies, capabilities and culture, and may have different impacts on an organization's performance (Li, Vanhaverbeke & Schoemakers, 2008).

Research on exploratory and exploitative innovation shows that different type of innovation activities lead to different performance, and the antecedents of both activities are different (Jansen et al., 2006; 2009). Firms that engage in exploratory innovation pursue new knowledge and develop new product and services for emerging customers or markets. Firms that engage in exploitative innovation build on existing knowledge and try to improve the quality and cost of existing product and services for existing customers (Benner & Tushman, 2003).

In international outsourcing, OEM manufacturers make tangible and intangible investments in equipment, operating procedures, and systems that are specialized to requirements of a particular buyer (Stump & Heide, 1996; Zaheer & Venkatraman, 1995). OEM manufactures provide manufacturing services according to OEM buyers' technical specifications or component performance requirements. These firms also design their manufacturing equipment and business processes for particular buyers in order to respond rapidly to their clients' demands. In order to acquire and keep the orders from particular buyers, OEM manufacturers will adopt exploitative innovation strategy to improve the product quality and reduce the cost of product. We thus propose:

Hypothesis 1*: The higher level of a firm's asset-specific investment, the more likely the OEM manufacturer will pursue exploitation innovation.*

Most of the companies in Taiwan IT industry are OEM manufacturers. To maintain close relationship with their OEM customers, these companies have to invest specific asset for their main OEM buyers, and the level of buyer's switching cost also influence the relationship and bargaining power between OEM manufacturer and their OEM buyers (Kaufman et al., 2000). OEM manufacturers will pursue exploitative innovation to improve the efficiency and quality of product to keep the orders from their OEM buyers, when the level of buyer's switching cost is higher.

Hypothesis 2*: The higher level of the OEM buyer's switching cost, the more likely the OEM manufacturer will pursue exploitation innovation.*

Researchers have founded that isolating mechanisms create barriers to impede competitors from imitating resources, capabilities and strategies (Grant, 1995; Mahoney and Pandian, 1992; Rumelt, 1984). Some isolating mechanisms are a mixture of resources characteristic, such as transparency, immobility, uncertain inimitability, low tradability, and durability whereas others involve managerial practice. As the isolation mechanisms holding by OEM manufactures are powerful, they can avoid imitation by competitor. Therefore, OEM manufacturers can adopt exploratory innovation to find out new market opportunities. We thus propose:

Hypothesis 3*: The more powerful of a firm's isolating mechanisms, the more likely the OEM manufacturer will pursue exploratory innovation.*

Both OEM and OBM strategy are important to OEM manufacturers for short-term and long-term profitability and viability. But the competences these two strategies needed are different. When OEM manufacturers engage in OEM business, they have to improve their production and process efficiency and the quality of their product to keep the orders from international OEM buyers. In contrast, when OEM manufacturers engage in OBM strategy, they have to develop their marketing ability and promote their product to emerging customers. Therefore, we posit that OEM manufacturers pursuing exploratory innovation can increase the production efficiency needed to maintain the competitive advantage of OEM business. In contrast, OEM manufacturers pursuing exploitative innovation can increase the marketing ability needed to create their own brand. We thus propose:

Hypothesis 4*: Exploitative innovation has a positive effect on the ability building of production efficiency.*

Hypothesis 5*: Exploratory innovation has a positive effect on the ability building of marketing.*

METHODOLOGY

Data and sample

Data for this study were obtained by questionnaire survey of electronic manufacturers in Taiwan. Our sample population consisted of 500 randomly selected electronic manufacturers listed in the directory of 2008 Taiwan computer manufacturers and suppliers published by CCIS. In the end, we received usable questionnaires from 103 firms, representing a response rate of 20.60%.

Measures

Exploratory and exploitative innovation. We define exploratory and exploitative innovation as follows. Exploratory innovation offers new designs, create new markets, and develop new channels of distribution. Conversely, exploitative innovation broadens existing knowledge and skills, improves established designs, and expands existing products and services, and increases the efficiency of production. Therefore, following He & Wong (2004), we measure these two types of innovation activities by eight Likert-scale items.

Competencies needed in OEM and OBM. We use the concept of “production efficiency” and “marketing ability” to represent the competencies needed in OEM and OBM. Based on previous research by Dyer(1996) and Liu et al.(2000), we use five items such as “production yield”, “process capability” and “speed of delivery” to measure manufacturing efficiency.

Asset-specific investment. Asset-specific investment means that OEM manufacturers make tangible and intangible investments in equipment, operating procedures, and systems that are specialized to requirements of a particular buyer. This variable was measured with 10 items developed by Kang, Mohaney & Tan(2008).

Buyer's switching cost. Switching cost is the perceived cost of changing suppliers by OEM buyer. Based on Kaufman et al. (2000), we use 9 items to measure buyer's switching cost.

Isolating mechanism. Isolating mechanisms are a mixture of resources characteristic, such as transparency, immobility, uncertain inimitability, low tradability, and durability. In this study, we use the concept of inimitability and non-substitutable to describe isolation mechanism. Following Mei-wei Wong (2001), we measure this variable by six Likert-scale items.

Control variables. In this study, we controlled for the industry sector, the age and size of OEM manufacturers.

RESULTS

Table 1 presents the hierarchical regression results between organizational antecedents and exploratory and exploitative innovation. According to this table, firm's asset-specific investment was significantly to exploitative innovation, which supported hypothesis 1. Hypothesis 2, which posited an association between OEM buyer's switching cost and exploitative innovation, was also supported. As for the association of a firm's isolation mechanisms with exploratory innovation, isolation mechanism was significant, which supported hypothesis 3.

Table 1

Hierarchical regression results: Effects of antecedents on exploratory and exploitative innovation

	Exploratory innovation		Exploitative innovation	
	Model 1	Model 2	Model 3	Model 4
Controls				
industry	-0.06	-0.11	-0.04	-0.04
Size	-0.02	-0.03	0.09	-0.12
Age	0.02	0.01	-0.22^{*}	0.05
Antecedents				
Asset-specific investment		0.15		0.28^{**}
Switching cost		-0.01		0.21^{**}
Isolation mechanisms		0.41^{**}		-0.13
R^2	0.03	0.20	0.05	0.30
ΔR^2		0.17^{**}		0.25^{**}

$^{*}p<0.05$ $^{**}p<0.01$

Table 2 shows the hierarchical regression results between exploratory and exploitative innovation, and the competences needed in OEM and OBM. The proposed positive relationship between exploitative innovation and production efficiency (the competences needed in OEM) was supported (hypothesis 4). Hypothesis 5, which posited a positive association between exploratory innovation and marketing ability, was also supported.

Table2 Hierarchical regression results: Effects of exploratory and exploitative innovation on competencies needed in OEM and OBM.

	Production efficiency		Marketing ability	
	Model 1	Model 2	Model 3	Model 4
Controls				
Sector	0.06	0.07	0.12	0.12
Size	0.22^{*}	0.19^{*}	0.20^{*}	0.17
Age	-0.11	-0.05	0.06	0.08
innovation				
Exploratory innovation		-0.20^{*}		0.45^{**}
Exploitative innovation		0.50^{**}		-0.19^{**}
R^2	0.06	0.27	0.04	0.21
ΔR^2		0.21^{**}		0.17^{**}

$^{*}p<0.05$ $^{**}p<0.01$

DISCUSSION AND CONCLUSION

The finding of this study revealed that exploratory innovation and exploitative innovation had different antecedents. When the level of investment on specific asset and buyer’s switching cost are higher, OEM manufacturers will tend to pursue exploitative innovation. But when OEM manufacturers hold a set of strong isolation mechanisms, they will pursue exploratory innovation. We also find evidenced that if OEM manufacturers hope to maintain the competitive advantage on its OEM business, they should pursue exploitative innovation to build competences needed in OEM business. But if OEM manufacturers hope to develop its OBM strategy, they have to pursue exploratory innovation. Expect the positive effect of exploratory and exploitative innovation on the competences needed in OEM and OBM respectively, we also find that exploitative innovation was negatively significant to the competences needed in OEM and exploratory innovation was negatively significant to the competences needed in OBM. Therefore, OEM manufactures should balance their exploitative and exploratory innovation activities carefully.

REFERENCES

References available upon request from the first author

THE STUDY OF SERVICE GAPs IN THE HOSPITAL

Shih-Wang Wu
Assistant Professor, Dept. of Hospital and Health Care Administration,
Chia Nan University of Pharmacy & Science, Taiwan
PhD Candidate, Graduate Institute of Business Administration,
National Chung Cheng University, Taiwan
+886-6-2664911 ext.5225, scottwu101@mail.chna.edu.tw

Hsien-Jui Chung
Professor, Dept. of Business Administration,
National Chung Cheng University, Taiwan
+886-5-2720411 ext.34313, hjchung@ccu.edu.tw

ABSTRACT

The main purpose of this research is to probe into the gap between physicians and their patients. The physicians are major service provider and the patients are their customers. As we know that customers always want more and they are always dissatisfied. The study want to know the gap in the hospital, thus administrators can do something to improve service quality and make customers more satisfied. One kind of the objects is doctor, and another is IPD patient in a medical center in Taiwan. The result of study shows that patients are almost unsatisfied because the major opinions are negative between patients' expectation and experiences actually. And then we adopt comparative analysis of the service attention degree of the doctors and their patients, we can know why the patient's real feeling and expectation to be mostly negative. In 11 items, there are four items significant of eight negative items. The hospital administrators and doctors should take serious and contact with their customers frequently in order to catch what is the key point that customers really want.

Keywords: service quality, patient's satisfaction, service gap, expectation, experience

INTRODUCTION

In recent years, consumers' right consciousness runs high gradually, and there are more and more medical lawsuits produced. Because there was particular, and quite specialized in medical industry in the past, so people do not know much about the treatment even though it is mal-medication. But people's relevant medical knowledge of the treatment also increases much more, and consumers will require best service quality. Department of Health in Taiwan adopted service quality as an indicator in the Hospital Accreditation since 1988, and hopes to improve the service quality in the hospital. Besides the request of government, every hospital starts to reconsiderate their service process and expects to satisfy their patients in order to get better public praise. Besides regarding fine medical quality in order to absorb patient go to institute to seek medical advice, the hospital manager gradually realizes the staff's service quality will influence the will that the people seek medical advice.

SERVICE QUALITY

Some scholars suppose that for customers or service providers, it is much more difficult to measure the service quality than other product. The conception of service quality is conducted by customers to make comparison with the results between the expected and actual perceived service. Also, the appraisal of service quality is not only based on service results but also include the appraisal about the process of service delivery. Namely, the service quality conceived by customers is determined by the difference between the expected and perceived service quality. Whenever the service expected by customers is higher than that perceived by customers, it means the service is unsatisfactory. Within the increasing difference, it shall finally cause the unsatisfied condition. Oppositely, whenever the perceived service is higher than that of expected, it means the customers are satisfied. With the increasing difference, it causes the inclination reaching the ideal service quality (Parasuraman, 1985) [1]. Gronroos (1990) also supposed that before customers purchase the service, they shall previously propose a certain degree of expected service. After they receive the service, the service experience is arisen. The difference between both exactly means the overall perceived quality. If the experienced quality cannot reach the expected quality, then the overall perceived quality is poor; otherwise, it is excellent quality [2].

Hospitals are certainly within a part of the said service industries. However, within the service industries, what category can hospitals fall into? Sasser (1978) divided the output of service industries into actual goods and intangible service. From results, there are roughly 20% of tangible goods and 80% of service available in hospitals. It indicates that the intangible output occupies very high proportion [3]. Berry &Bendapudi (2007) note that medical services differ from other services because the customers may be ill or stressed and the service–which is need-based rather than want-based–may itself involve risk(e.g., hospital acquired infections; medical treatment errors) [4].

Elston et al.(2008) according to research Major US corporations and consumer groups are demanding more accountability for their health care expenditures. In response, the federal government, specialty boards, and state medical boards are evaluating ways to implement objective measures of quality. Many dermatologists already choose to participate in quality measurement and improvement activities [5]. Lovelock (1983) integrated the opinions of Sasser and also added some new categorizing variables. The said categorizing method separately includes basic demand characteristics, service content and benefits, and service delivery procedures [6]. Under the said categorization, the category of hospital service means the direct tangible service action available for people. To sum up the aforesaid categorizations proposed by scholars, we can know that hospitals are the service under high frequency contact and they belong to the service providers with more intangible service items than tangible items.

Lee & Yom (2007) showed their study results that nurses’ expectations and performance were higher than those of patients, while patients’ overall satisfaction with nursing and medical care was higher than that of nurses. There was a strong positive relationship between satisfaction with nursing and medical care and intent to revisit the hospital for both groups. The performance was relatively lower than expectations, resulting in poor nursing care quality. Differences between expectations and performance for both patients and nurses need to be further reduced [7]. For patients' conception toward service levels and service quality, it is a type of attitude representation. Thus, we can measure by using the model of attitude measurement. About the psychology and marketing science, attitude is exactly the most subject for measurement. Also, there are numerous models proposed by scholars. After making the suitability comparison, we adopt the SERVPERF Scale proposed by Cronin (1992) [8]. The said is easy for operation and it is also well testified with the effective validity to measure customers' appraisal toward service quality (Cronin, 1994) [9]. The representation way is shown as below. Wherein, Qo means customers' overall attitude toward service quality; Pi means the

customers' actual perception about service quality and Ii means customers' emphasis degrees about service quality.
The equation is as follows:

$$Qo = \sum_{i=1}^{n} (Ii \times Pi)$$

METHODOLOGY

According to the above-mentioned literature review, after consulting and synthesizing the research of each scholar, sum up several more important variables and information wanted to collect. The main purpose of this research is to probe into the gap between physicians and their patients. So we use the questionnaire to collect their attitude, feeling and demographic variables of doctors and patients. We adopt closed questionnaire, and use positive description to measure with Likert's scale. We choose a medical center in Taiwan to be our sample by "convenient sampling". We sample 50 doctors randomly in among all physicians who have inpatients, and we make random sampling six inpatients per doctor, and we will collect 300 inpatients altogether finally. The execution time of research questionnaire was divided to two parts. The first half a period of time, it is the execution of doctor's questionnaire, but the latter half section is the sample investigation of the inpatients. Such as meeting the persons who are less than six inpatients of doctor, we sampled separately in another day until six. If the patient is the serious illness, or pediatrics patient, questionnaire is transferred to relative or parents.

RESULT

We get quite high Cronbach. In doctor's questionnaire, the Cronbach α value of "attention of service quality" is 0.8855. In patient's questionnaire, the Cronbach α value of "attention of service quality" and "perception of service quality" is 0.8732 and 0.9288. About the difference analysis of patients' emphasis degree and actual perception, most gaps belong to negative. Namely, within most items, patents' expectations are higher than that of actual perception and they are also featured with significant difference. However, only two of the items show no significant difference, namely physicians' punctual ward patrol and ask patients for ways of treatment. It also means the physicians can reach patients' expectation on 2 items. There are only 2 items showing positive but no significant difference, and this situation also means the service have not reach patents' expectation yet. See table 1.

Following that, we continue to explore the emphasis difference about service quality at both physicians and patients. Among the groups with significant difference, there are 2 items showing that physicians' emphasis is higher than patients'. It means physicians really care about patients' syndrome and respect patients' privacy. In view of physicians' stance, what they attempt most is to cure patents'. Thus, syndrome tracking shall be the points receiving much more attention from physicians. About the respect to privacy, because people always are unwilling to invade others' privacy, we can find physicians' special prudence on the said pints when they are practicing medicine. Based on aforesaid factors, physicians naturally show higher emphasis than that of patients. About others, they include the items like immediate reach of attending physicians on demands, punctual ward patrol, detailed description for patients' kindly and detailed description for diagnosis and prescription. Patents' emphasis is higher than that of physicians. It is possibly because physicians only show their emphasis on the item, namely immediately reachable access to attending physicians.

However, the said items show no critical influence on medical treatment; thus, it really exists in significant difference from patients' conception. Within the items without significant difference, almost all the average values of inpatients are rated higher than physicians. It is possibly the reason why most patient gaps belong to negative gaps. See table 2.

DISCUSSION AND CONCLUSION

It is required to focus the exploration about the perception gaps from patients and it is quite important. We shall communicate with physicians and know about the ideas from the public so that it is available to enhance patients' satisfaction. Because the public show higher knowledge levels now, they naturally request service with higher standard. Thus, if there are numerous gaps existing as the research results, the situation is quite risky. Because, under the said situation, if there are better selections available for patients, our hospital will lose customers, For a long time, it shall definitely influence the management status of hospital. It is required to discuss with physicians and find the way to enhance patients' satisfaction and know about patients' emphasis. We can inspire physicians to change their traditional notions and show increasingly attentive care about patients' conception.

We shall make better allocation for physicians' reward and service quantity. If we establish the reasonable PF systems, the more patients can be served and the higher salary shall be paid. Thus, in consideration of pay income, physicians naturally create the close relation with patients to enhance service quality and oral admiration so that the old patients shall be well kept and new patients can be solicited. As such, the hospital income is also increased and the 3-win happy ending of the said three parties can be reachable.

Table 1 Comparative analysis of emphasis degree and actual perception of patients

Items	Classification	Average score	P value
The physician is considerate and easily to communicate	emphasis degree actual perception	4.6233 4.3600	0.000
The doctor cares about the patient's condition	emphasis degree actual perception	4.7600 4.4400	0.000
Patient can find the doctor immediately when necessary	emphasis degree actual perception	4.5667 3.9967	0.000
The doctor enriches clinical experience and enough professional ability	emphasis degree actual perception	4.8067 4.4467	0.000
The doctor respects the patient's right of privacy	emphasis degree actual perception	4.0367 4.0733	0.508
The doctor is intimate and courteous	emphasis degree actual perception	4.4133 4.4567	0.300
The doctor listens to the patient telling patiently about the condition	emphasis degree actual perception	4.5433 4.3900	0.001
The doctor patrols the room on time	emphasis degree actual perception	4.2500 4.1767	0.163
The doctor explain for the patient patiently about the disease	emphasis degree actual perception	4.5667 4.3733	0.000
Prudent, careful when the doctor makes a diagnosis	emphasis degree actual perception	4.7133 4.4233	0.000

Items	Classification	Average score	P value
The doctor seek the suggestion of asking the patient while determining the way of treating	emphasis degree actual perception	4.2733 4.1967	0.170
The doctor states the diagnosis made and open prescription carefully to the patient	emphasis degree actual perception	4.5133 4.2700	0.000

Table 2 Comparative analysis of attention to service quality between patient and doctor

Items	Classification	Importance	The difference of average	P value
The doctor cares about the patient's condition	Doctor Patient	4.9000 4.7600	0.1400	0.007
Patient can find the doctor immediately when necessary	Doctor Patient	4.1200 4.5667	-0.4467	0.001
The doctor enriches clinical experience and enough professional ability	Doctor Patient	4.7600 4.8067	-0.0467	0.513
The doctor respects the patient's right of privacy	Doctor Patient	4.5800 4.0367	0.5433	0.000
The doctor is intimate and courteous	Doctor Patient	4.5600 4.4133	0.1467	0.072
The doctor listens to the patient telling patiently about the condition	Doctor Patient	4.4200 4.5433	-0.1233	0.192
The doctor patrols the room on time	Doctor Patient	3.8000 4.2500	-0.4500	0.000
The doctor explain for the patient patiently about the disease	Doctor Patient	4.3400 4.5667	-0.2267	0.014
Prudent, careful when the doctor makes a diagnosis	Doctor Patient	4.6200 4.7133	-0.0933	0.238
The doctor seek the suggestion of asking the patient while determining the way of treating	Doctor Patient	4.2600 4.2733	-0.0133	0.918
The doctor states the diagnosis made and open prescription carefully to the patient	Doctor Patient	4.2400 4.5133	-0.2733	0.013

REFERENCES

[1] Parasuraman, A., Zeithaml, V. A., & Berry, L. L., *A conceptual model of service quality and it's implications for future research*, Journal of Marketing, Vol.49, fall, pp41-50, 1985

[2] Gronroos, Christian, *Service Management and Marketing*. Lexington, pp41, 1990

[3] Sasser, W. Earl, R. Paul Olsen, & D. Daryl Wyckoff, *Management of service operations – Text, cases, and readings*, Allyn and Bacon, 1978

[4] Berry L.L., & Bendapudi, N., *Health care: a fertile field for service research*. J Serv Res, Vol.10, No.2, pp11-22. 2007

[5] Elston, D.M., Sieck, C.K., Sullivan, J.N., Behal, R., Kaleba, E.O. *Developing ambulatory care physician performance measures*. JAMACAD DERMATOL(SEPTEMBER), pp505-513, 2008

[6] Lovelock, C. H., *Classifying services to gain strategic marketing insights*, *Journal of Marketing*, Vol.47, Summer, pp9-20, 1983

[7] Lee, M.A. & Yom, Y.H., *A comparative study of patients' and nurses' perceptions of the quality of nursing services, atisfaction and intent to revisit the hospital: A questionnaire survey.* International Journal of Nursing Studies 44, pp545–555, 2007

[8] Cronin, J. Joseph & Taylor, Steven A., *Measuring Service Quality ∶ A reexamination and extension*, *Journal of Marketing*, Vol.56, July, pp55-68, 1992

[9] Cronin, J. Joseph & Taylor, Steven A., *SERVPERF vs. SERVQUAL ∶ Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality*, *Journal of Marketing*, Vol. 58, Jan,pp125-131, 1994

The Determinants of Consumers' Satisfaction with Hotels in China

Jiani Jiang, Roger Williams University, Bristol, RI
Miao Zhao, Roger Williams University, Bristol, RI

INTRODUCTION

China has become one of the leading travel destinations in the world. The World Tourism Organization has projected that China will become the world's top tourism destination by 2014 (China Daily 2005). Tourism satisfaction with their hotels will significantly influence their travel experience and satisfaction with travel destinations. Thus, this research examines how hotel attributes – price, room cleanliness, service, facilities, hotel location, room comfort, and stay value – affect guests' satisfaction with hotels. Star ratings of hotels are mainly determined by physical facility quality and service quality (Ingram and Roberts 2000). Therefore, customers use hotel star ratings as guidelines when they select hotels. This research tends to examine the different impacts of hotel attributes on satisfaction across hotels with different star-ratings. With China's economic development, more and more international hotel brands have entered China. This research also tries to explore the different determinants of customer satisfaction between international hotels and domestic hotels.

Different from previous studies, this study examines the data compiled directly from www.elong.com, the leading online travel planning and hotel booking website in China. The website allows hotel guests to evaluate their stay at hotels in terms of the price, room cleanliness, hotel service, hotel facilities, hotel location, room comfort, stay value and satisfaction with hotels.

LITERATURE REVIEW

China has surpassed all other Asian countries and become the leader in international tourist arrivals and receipts in the region over the last ten years (Ye 2008). The total revenue of China's tourism industry will increase from US$ 67.3 billion in 2002 (China Knowledge Press 2004) to the projected US$ 85 Billion in 2011 (RNCOS 2007). Tourists' satisfaction with hotels is one of the key elements which affect tourists' selection of and their satisfaction with travel destinations. Some researchers have examined the hotel attributes which affect hotel selection and satisfaction. For example, Callan (1996) identified that cleanliness as the most important factor among the 166 hotel attributes including location, image, price/value, competence, access, security and so on. Dolnicar and Otter (2003) extracted 173 attributes such as previous experiences, recommendation, external rating of a hotel, location, price and so on that determined occupancy rates by reviewing 21 journals undertaken between 1984 and 2000. The authors reviewed empirical studies of the importance of hotel attributes and provided attributes rankings. Locker (2005), using data collected from focus groups, revealed that the availability of a room, the price and cleanliness were the three key attributes in the hotel room selection and satisfaction.

Some other researchers focused on the different views from the hotel guests and managers. For example, Locker (2003) reported that both of the business guests and accommodation managers believed that the cleanliness of the hotel affect the accommodation selection most. However, the business guests rated the "bathroom and shower quality" and "comfort of mattress and pillow" higher than other attributes. The managers convinced "courteous, polite, well- managed staff", "enthusiasm, and commitment of staff", and "efficiency of front desk" are more important.

More recently, researchers have started to concentrate on Chinese hotel industry. Yu and Gu

(2005) argued that after two-decade of rapid development, Chinese hotel products were diversified and with high quality. Ryan and Gu (2007) surveyed 941 hotels guests who had recently stayed in a Chinese hotel six months prior and found that the key determinants of satisfaction were the external environment, star-rating and cleanliness of the bedroom. This research tried to contribute to our understanding of Chinese hotel industry by answering the following question—which attribute influence the guests' satisfaction with hotels in China?

China developed its five-star hotel ranking system in the early 1990s to standardize its fledgling hotel operations and service (Yu 1992). The introduction of hotel ranking system enhances the hotel management and service standards of hotels and to protect the interests of hoteliers, travel companies, and customers. The China National Tourism Administration (CNTA) issued "Regulations on Star standard and Star-Rating of Hotels of the PRC" to regulate service quality and facilities and to enable China's hotels to meet the international standards (Liu 2006). Given the different requirements for hotels with different star ratings, the second question this research intends to answer is how hotel attributes impact satisfaction differently across hotels with different star ratings. Since China's reform and opening up, China has become an increasingly attractive destination for the international hotels; at the same time, domestic hotels has grown rapidly. The competition between international hotels and domestic hotels has become increasingly intense during the past decades. On the one hand, international hotels have good reputation and wealthy experience. On the other hand, domestic hotels have more understanding of local tourists. Thus, the third research question is whether the impacts of hotel attributes on satisfaction are different between international hotels and domestic hotels.

EMPIRICAL INVERTIGATION

An empirical study was designed to answer the above mentioned three research questions. Secondary data were compiled directly from www.elong.com which collects and reports real customer ratings of hotels in China. After staying at a hotel booked at elong.com, each consumer is asked to rate the hotel's room price, room cleanliness, service, facilities, hotel location, comfort, stay value, and the overall satisfaction. Each attribute except price is rated on a 0-5 scale which from "very dissatisfied" to "very satisfied". Online hotel ratings are publicly reported when consumers have rated this hotel. Each rating is the average of evaluations from all respondents. Even though there are some limitations to the elong data such as not all the guests choose to rate the hotels after their stay and not all the hotel attributes discussed in the existing literature are selected by elong.com, the major advantages of the elong data suggest they are quite useful for the current research. The data are collected from real customers staying at hotels which assume high external validity of the data. Moreover, a great amount of customers responding to the survey after an actual stay suggests the credibility of the data.

Data Collection

Data was collected from September 2009 to October 2009. Ratings for 785 hotels in Beijing were collected as available data for the time period. Beijing was chosen because it is the most popular city for traveling in China. As the capital city and the political, educational and cultural center of China, Beijing is rich in historical sites and important government and cultural institutions.

ANALYSIS AND RESULTS

The first analysis focused on the effects of individual attributes on guests' satisfaction with hotels. All the hotel attributes were regressed on satisfaction (Table 1). The two most important attributes were "Comfort" and "Facilities", followed by "Location", "Service", "Cleanliness", and "Value". "Price" did not significantly contribute to the evaluation of guests' overall satisfaction with hotels.

Table 1: Effects of Hotel Attributes on Customer Satisfaction

Attribute	Beta	p
Price	-.004	.541
Cleanliness	.184	.000
Service	.189	.000
Facilities	.208	.000
Value	.183	.000
Comfort	.225	.000
Location	.195	.000

R^2 =.981

A set of ANOVAs were conducted to examine whether guests evaluate hotels with different star ratings differently (Table 2). All ANOVA tests were statistically significant, which suggested that hotel guests evaluated different star-rating hotels differently. Tukey's HSD tests were then conducted and the results indicated that there were no significant differences among one-, two- and three-star hotels where five-star hotels were perceived better on all attributes than other hotels. More specifically,

Table 2: Comparison of Attribute Ratings across Different Star Ratings

Star	Budget (n=216)	One (n=28)	Two (n=40)	Three (n=176)	Four (n=215)	Five (n=110)	F	P
Price	205.23[a](43.35[b])	175.32(42.99)	202.83(39.79)	277.05(91.52)	418.95(147.68)	820.20(432.30)	185.52	.000
Cleanliness	3.78(.48)	3.45(.36)	3.45(.56)	3.61(.54)	3.86(.43)	4.17(.39)	27.11	.000
Service	3.71(.49)	3.50(.36)	3.46(.49)	3.62(.58)	3.73(.48)	3.98(.52)	10.15	.000
Facilities	3.44 (.41)	3.13(.38)	3.21(.53)	3.30(.59)	3.58(.50)	4.03(.45)	38.98	.000
Location	3.61(.47)	3.45(.35)	3.48(.36)	3.48(.58)	3.58(.47)	3.79(.51)	4.66	.000
Comfort	3.51(.44)	3.30(.35)	3.32(.62)	3.44(.57)	3.71(.46)	4.07(.45)	32.72	.000
Value	3.63(.42)	3.67(.32)	3.63(.62)	3.74(.57)	3.77(.47)	3.89(.40)	6.55	.000
Satisfaction	3.62(.37)	3.42(.30)	3.43(.45)	3.54(.50)	3.71(.38)	3.99(.35)	23.29	.000

[a:] mean
[b:] standard deviation

- "Price" of budget hotels and two-star hotels were the same. It was lower than three-star, four-star and five-star hotels, but higher than one-star hotels. The five-star hotels had the highest price.
- Budget hotels had higher ratings than one-star, two-star and three-star hotels on "Cleanliness", and had similar rating as four-star hotels, lower rating than five-star hotels.
- Customers rated the highest on "Service" for five-star hotels. Four-star hotels had the second highest rating; budget hotels had the third, while one-, two-, and three-star hotels got the lowest rating.
- Budget hotels had higher rating than one-star, two-star and three-star hotels on "Facilities", but the rating was lower than four-star and five-star hotels.
- Budget hotels had higher rating on "Location" than one-star, two-star, three-star and four-star hotels, while the rating was lower than five star hotels.
- Budget hotels had a higher rating than one-star, two-star and three-star hotels on room "Comfort",

but lower than four-star and five-star hotels.

- Five-star hotels were rated highest on hotel “Value”, and all other hotels had the similar rating.
- As to “Overall Satisfaction”, budget hotels had slight lower rating than four star hotels, but higher than one-star, two-star and three-star hotels. Five star hotels were rated the highest.

The next set of analyses focused on the effects of individual attributes on guest overall satisfaction on each star-rating-hotel (Table 3). Since there were no significant differences found among one-, two-, and three-star hotels, we combined them into one category. All the regressions were statistically significant.

- Budget hotels: all the attributes affected the customers’ overall satisfaction. “Value” influenced the most, “Comfort” was the second, “Service” and “Location” were more important than “Cleanliness” and “Facilities”, and “Price” affected the least.
- One-, two-, and three-star hotels: “Facilities” and “Comfort”, contributed the same to the customer satisfaction. “Cleanliness” and “Value” were more important than “Location” and “Service”. “Price” was not a factor affecting satisfaction.
- Four Star hotels: “Value” and “Comfort” contributed the most positively to customer satisfaction. “Facilities” and “Service” were more important than “Location” and “Cleanliness”. “Price” was not related.
- Five Star hotels: “Service” affected the satisfaction most; “Location”, “Facilities” and “Comfort” were more important to the satisfaction than “Cleanliness” and “Value”. “Price” didn’t significantly contribute to Satisfaction.

Table 3: Attributes’ Impacts on Satisfaction across Different Star Ratings

Star	Budget	One/Two/Three	Four	Five
Price	.023[a]	.014	.001	-.029
Cleanliness	.190[a]	.199[a]	.178[a]	.179[a]
Service	.207[a]	.160[a]	.209[a]	.256[a]
Facilities	.128[a]	.210[a]	.213[a]	.237[a]
Location	.205[a]	.161[a]	.193[a]	.248[a]
Comfort	.227[a]	.210[a]	.220[a]	.211[a]
Value	.230[a]	.193[a]	.223[a]	.173[a]
R^2	.981[a]	.968[a]	.990[a]	.991[a]

[a] <.05

Table 4: Comparison of Attribute Ratings between International and Domestic Hotels

Attributes	International	Domestic	T	P
Price	617.73	336.95	80.551	.000
Cleanliness	4.03	3.76	19.323	.000
Service	3.86	3.69	7.297	.007
Facilities	3.78	3.47	23.218	.000
Location	3.69	3.57	4.174	.041
Comfort	3.88	3.59	21.550	.000
Value	3.84	3.72	4.340	.038
Satisfaction	3.85	3.64	17.083	.000

a: mean

To compare hotel guests' evaluations between international and domestic hotels, two groups were formed – international brand hotels and domestic hotels, and t-tests were conducted (Table 4). The price of international hotels was much higher than domestic hotels'. And international hotels were rated higher on all attributes including overall satisfaction than domestic hotels by guests.

The following set of analyses examined whether impacts of hotel attributes on customers' overall satisfaction were different between international hotels and domestic hotels. Two regressions were run for international hotels and domestic hotels, respectively (Table 5). Both regressions were statistically significant. All the attributes affected the customers' overall satisfaction with international hotels except "Price". "Facilities" influenced satisfaction the most. "Comfort" and "Service" were more important than "Location" and "Value". Lastly, "Cleanliness" influenced the guests' satisfaction slightly.

For domestic hotels, "Comfort" influenced the most on the satisfaction. Moreover, "Value", "Service", "Cleanliness" and "Location" had the similar importance to customers' satisfaction. Finally, Price" was not significantly contribute to the evaluation of guests' overall satisfaction on domestic hotels.

Table 5: Attributes' Influence on Satisfaction between International and Domestic Hotels

Attributes	International	P^i	Domestic	P^d
Price	-.026	.302	.002	.781
Cleanliness	.140	.000	.188	.000
Service	.219	.000	.189	.000
Facilities	.258	.000	.203	.000
Location	.197	.000	.183	.000
Comfort	.249	.000	.223	.000
Value	.182	.000	.197	.000

Discussion and Conclusion

With the development of Chinese hotel industry, we can also expect greater research on hotel attributes and the impacts of attributes on customer satisfaction. Our literature review had suggested that there is no consensus regarding hotel attributes and how they affect customer satisfaction. Based on the analysis of the elong data collected from real hotel guest ratings, we can draw some conclusions regarding the influences of hotel attributes on satisfaction with hotels in Beijing.

- First of all, our analysis found that "Comfort" is the most important attribute affecting customer satisfaction. "Facilities" is next in importance. "Value" and "Service" are also important. Hotels should focus on improving "Comfort" and "Facilities" to be more competitive.
- Five-star hotels were rated the highest on all the attributes including the customer satisfaction. Five-star hotels target at the high income consumers. For these people, they do not concern much about price; they pursue a joy of high quality of facilities and service and the convenience to reach the attractions and downtown area when they stay at hotels. Our analysis found that the price of budget hotels was higher than one-star and two-star hotels. The customer satisfaction of budget hotels was also higher than one-star and two-star hotels, and sometimes even higher than three-star hotels. In China, international budget hotels have a great market share. These budget hotels, majorly owned by American companies, follow the same standard in China as in U.S.A. Later when the domestic budget hotels emerge; they follow the international concept and standard. Therefore, the budget hotels have comparably higher ratings on all hotel attributes. The budget hotels in China target at the growing business travelers and tourists who want economical but comfortable

accommodation in cities (China Daily 2004).

- The analysis also suggested that international hotels had higher ratings than domestic hotels. As China's hotel industry prepares for international challenges ahead, it is critical for its operators to be aware of the weaknesses of domestic hotel operation (Yu and Gu 2005). China's hotel industry prospers after its opening up; however, the hotel operators do not have much experience on how to develop and manage hotels at a high standard. Meanwhile, the hotel industry in developed countries is already mature, and they have a great amount of human resource on managing hotels. Therefore, domestic hotels should learn from international hotels. As this research identified, enhancing "Comfort" and "Facilities" is the most pressing matter at this moment.

Reference:

Callan, RJ. 1996, *An appraisement of UK business travelers' perceptions of important hotel attributes*, Hospitality Research Journal, 19(4), pp. 113-27.

China Daily. 2005, *China Heading for Top Spot in World Tourism.* http://www.chinadaily.com.cn/china/2007-07/02/content_907643.htm

China Daily. 2004, *Budget Hotels Gain Popularity* http://www.chinadaily.com.cn/english/doc/2004-06/24/content_342145.htm

China Knowledge Press. 2004, *China Tourism Industry.* http://www.researchandmarkets.com/reportinfo.asp?report_id=220246

Dolnicar, S., & Otter, T. 2003, *Which hotel attributes matter? A review of previous and a framework for further research*. In Asia Pacific Tourism Association 9th annual conference, ed. T. Griffin and R. Harris. Sydney, Australia: University of Technology Sydney.

Ingram, P. & Roberts, P. 2004, *Friendships among Competitors in the Sydney Hotel Industry*. American Journal of Sociology. 106(2), 387-423.

Liu, Z. 2006, *Tourism and Hotel Development in China: From Political to Economic Success.* Journal of Vacation Marketing, 12(2), 191-192.

Lockyer,T. 2002, *Business guests' accommodation selection: The view from both sides*. International Journal of Contemporary Hospitality Management, 14(6), 294-300.

Lockyer, T. 2005, *Understanding the dynamics of the hotel accommodation purchase decision*. International Journal of Contemporary Hospitality Management, 17(6/7), 481-492.

RNCOS. 2007, *China Tourism Industry Analysis 2007-2011*. http://www.researchandmarkets.com/reportinfo.asp?report_id=575245

Ryan, C. & Gu, H. 2007, *Perceptions of Chinese Hotels*. Cornell Hotel and Restaurant Administration Quarterly, 48(4), 380-391,357.

Ye, L. 2008, *Tourism and China's Development: Policies, Regional Economic Growth and Ecotourism/Tourism in China/Tourism and Hotel Development in China: From Political to Economic Success.* Journal of the American Planning Association, 74(1), 150-151.

Yu, L. 1992, *Seeing Stars: China's Hotel Ranking System*. Cornell Hotel and Restaurant Administration Quarterly, 33(5), 24-27

Yu, L.& Gu, H. 2005, *Hotel Reform in China: A SWOT Analysis.* Cornell Hotel and Restaurant Administration Quarterly, 46(2), 153-169.

EMPLOYEE OPINION EXPRESSION AND SUPERVISORY PERFORMANCE EVALUATION

Hsin-Hua Hsiung, National Dong Hwa University, (+886) 3-863-3027, hsiung@mail.ndhu.edu.tw

ABSTRACT

This study examined the relationship between employee opinion expression and supervisory performance evaluation. Our review of the literature indicated that opinion expression behaviors included at least three different forms: voice, dissent, and cynicism. Self-rating and other-rating were both used to measure each form of opinion expression behaviors. Data were collected from a sample of 270 employee-supervisor pairs in Taiwan. Results show that voice and dissent are positively related to supervisory performance evaluation, but cynicism is negatively related to supervisory performance evaluation.

Key words: voice, dissent, cynicism, performance evaluation

INTRODUCTION

Employees' opinions and ideas are important driving forces for organizational innovation and improvement (Zhou & George, 2001). Listening to the "truth" is a necessary capability for managers to make good or correct decisions (Collins, 2002). Management literature has repeatedly emphasized the importance of building open communication channels within organizational hierarchy. While most managers consider themselves to be open-minded and are willing to accept various opinions and ideas, many employees report that their organizations or supervisors do not encourage communication nor the sharing of information and knowledge (Beer & Noria, 2000). It reflects a phenomenon that employees do not trust in their supervisors' tolerance or magnanimity. Employees who exercise voice may worry about the risks to their reputation, interpersonal relationship, and/or career development (Klaas & DeNise, 1989; Krefting & Powers, 1998).

Many studies have endeavored to identify the antecedent variables of opinion expression (e.g. Janssen, Vries, & Cozijnsen, 1998; Tangirala & Ramanujam, 2008), but very few studies have focused on the personal consequences of opinion expression, which are exactly employees' significant concerns in mind. Therefore, the objective of this research is to examine whether employees' opinion expression would lead to advantageous or disadvantageous effects on supervisory performance evaluation.

LITERATURE REVIEW AND HYPOTHESES

Literature has indicated a variety of forms of opinion expression behaviors, including voice, dissent, and cynicism. Since managers may have different feelings for different opinion expression behaviors, the present study examines the relationship between each form of opinion expression behavior and supervisory performance evaluation respectively.

In theory and research, employee voice is a communicative behavior primarily from employee to supervisor. However, it can include "submitting and discussing suggestions and ideas about the subordinate's own functioning, about others and their problems, about organizational policies and practices, or about what needs to be done and how it could be done" (Janssen, Vries, & Cozijnsen, 1998). Many researchers treated voice as a form of organizational citizenship behavior, which reflects an employee's organizational participation and involvement (Podsakoff, MacKenzie, Paine, & Bachrach, 2000). Since voice is a promotive or pro-social behavior intending to improve the status quo rather than merely criticize (Van Dyne & LePine, 1998), such a form of opinion expression is more likely to be accepted and appreciated by managers than other forms of opinion expression. Voice behaviors usually play a constructive role in organizational change and development, so these behaviors are expected to contribute to organizational functioning and effectiveness. Thus, we predict that voice would be positively associated with supervisory rating of employee performance.

Hypothesis 1: *Employee voice is positively related to the performance evaluation that supervisors give to employees.*

Dissent is "the effort by individuals in the workplace to protest and/or to change the organizational status quo because of their conscientious objection to current policy or practice" (Graham, 1986). One feature of dissent is that an individual feels that he/she needs to stand apart from or against an organization. A dissenting individual would be dramatically fore-grounded and the outspokenness would make him/her distinct from so many conformable employees (Shahinpoor & Matt, 2007). While a dissenter provides conscientious reflection and important critical thinking to the organization, his/her acts are often unanticipated to the managers. Despite the potential value of dissent, most managers fail to recognize or appreciate this act, and some managers even try to oppress or silence it (Shahinpoor & Matt, 2007). Since dissenters take a stand in opposition to a "current policy or practice", in many cases, managers consider them troublemakers. Therefore, we predict that employee dissent would have negative effects on supervisory performance evaluation.

Hypothesis 2: *Employee dissent is negatively related to the performance evaluation that supervisors give to employees.*

Cynicism is usually defined as a negative response or attitude toward one's employing organization (Naus, Iterson, & Roe, 2007). Such a response or attitude comprises three dimensions: (1) a belief that the organization lacks integrity; (2) negative affect toward the organization; and (3) tendencies to disparaging and critical behavior toward the organization that are consistent with these beliefs and affects (Dean et al., 1998). Past literature has named many examples of organizational characteristics, practices, or events, such as managerial incompetence (Stanley, Meyer, & Topolnytsky, 2005), the absence of meaning in work (Cartwright & Holmes, 1006), a lack of sincere participation in decision-making processes (Fleming, 2005), and the deficient quality of leader-member exchange (Bommer, Rich, & Rubim, 2005; Cole, Bruch, & Vogel, 2006) that may lead to cynical experiences. Since the negative feelings associated with cynicism may undermine an employee's enthusiasm, creativity, and involvement in his/her work, we expect that employees' cynicism would have a disadvantageous effect on performance rating.

Hypothesis 3: *Employee cynicism is negatively related to the performance evaluation that supervisors give to employees.*

METHOD

Sample and Procedure

This study collected paired samples from administrative personnel and their direct supervisors in 56 junior high schools in Taiwan. In this study, employees' performance evaluations were provided by their supervisors. To measure employees' opinion expression behaviors (including voice, dissent, and cynicism), self-rating data (provided by employee themselves) and other-rating data (provided by supervisors) were both collected to avoid the problem of common method variance. This approach was also helpful for comprehending the perceptions of behavior performers and behavior observers.

A total of 422 paired questionnaires were distributed through mail delivery and personal interviews. Three hundred and eleven subordinates and 304 supervisors returned the surveys, making for a response rate of 74% and 72%, respectively. Eliminating questionnaires that could not be matched and those with missing data produced a final sample of 270 supervisor-subordinate pairs. Of the subordinates, 64.1% were women, with an average age of 38.3 years and average organizational tenure of 7.9 years. Of the supervisors, 44.1% were women, with an average age of 43.7 years and average organizational tenure of 9.4 years.

Measures

The questionnaires were in Chinese but three opinion expression scales—voice, dissent, and cynicism—were originally constructed in English. A translation and back-translation procedure were employed to ensure literal correspondence (Brislin, 1980). All of the independent and dependent variables were measured with a seven-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree).

Voice. An abbreviated version scale developed by Hagedoorn, van Yperen, van de Vliert, and Buunk (1999) was used to measure voice. This scale is comprised of five items, which have been used by Naus, van Iterson, and Roe (2007). Sample items are, “Try to work out solutions the organizations might benefit from,” and “Discuss the problem with your superior and try to work out a solution together.” The coefficient alpha of this scale is .90 for self-rating and .92 for other-rating.

Dissent. A two-item scale developed by Dooley and Frywell (1999) was used to measure dissent. Items are, “To what extent did you openly express a difference of opinion,” and “To what extent did you voice dissent while making the decision?” The coefficient alpha of this scale is .88 for self-rating and .84 for other-rating.

Cynicism. A six-item scale developed by Naus et al. (2007) was used to measure cynicism. Sample items are, “Withhold suggestions for improvements, because you think nothing is going to change anyway,” and “Shrug your shoulders at what management requires you to do.” The coefficient alpha of this scale is .62 for self-rating and .73 for other-rating.

Job performance rating. A four-item scale developed by Farh and Cheng (1997) was used to measure employees’ job performance. The ratings of this scale were provided by supervisors. A sample item is, “This employee makes significant contribution to the overall performance of our work unit.” The coefficient alpha of this scale is .95.

RESULTS

Table 1 presents the means, standard deviations, correlations, and reliabilities of all variables. Our results indicate that there is a significant correlation between employees’ perceptions and supervisors’ perceptions. The Pearson correlation coefficient between self-rating voice and other-rating voice is .26 ($p < .01$), the correlation coefficient between self-rating dissent and other-rating dissent is .26 ($p < .01$); the correlation coefficient between self-rating cynicism and other-rating cynicism is .27 ($p < .91$). Hence, these results suggest that employees’ perceptions and supervisors’ perceptions have a certain degree of consensus. Although employees may not express their cynicism in their supervisors’ presence, their supervisors are also aware of these cynical behaviors.

To examine the construct validity of the three forms of expression behaviors (including voice, dissent, and cynicism), we conducted confirmatory factor analyses (CFAs) for self-rating data and other-rating data respectively. With regard to self-rating data, the hypothesized three-factor model has a good fit to the data (GFI = .90, CFI = .91; IFI = .91, NFI = .87, RMSEA = .10). As far as other-rating data are concerned, the fit indexes also reveal a good model-to-data fit (GFI = .89, CFI = .92; IFI = .92, NFI = .90, RMSEA = .10). Furthermore, against the baseline three-factor model, several alternative nested models were also tested. Model comparisons reveal that the three-factor model fits better than any of the alternative nested model. These CFA results demonstrate that the constructs in this study have satisfactory discriminant validity.

The hypotheses were examined using the full model approach in structural equation modeling (SEM). The results from SEM analysis show that self-rating voice and other-rating voice are both positively related to supervisory performance evaluation ($\gamma = .18$, $p < .05$; $\gamma = .27$, $p < .05$), providing support for Hypothesis 1. The results related to self-rating dissent and other-rating dissent are inconsistent. While self-rating dissent has no significant relationship with supervisory performance evaluation ($\gamma = .07$, $p > .05$), other-rating dissent has a positive relationship with supervisory performance evaluation ($\gamma = .13$, $p < .05$). Since neither of the results is in line with our expectation, thus Hypothesis 2 is not supported. Finally, self-rating cynicism and other-rating cynicism are both negatively related to supervisory performance evaluation ($\gamma = -.14$, $p = .09$; $\gamma = -.65$, $p < .09$). These results provide support for Hypothesis 3.

DISCUSSION

Among the three forms of opinion expression behaviors, voice is usually treated as a typical organizational citizenship behavior in literature. Our results indicate that supervisors approve employees' voice behaviors and are willing to give positive feedback to the employees who have made constructive suggestions. While employees tend to think that supervisors may not accept their dissent or opposite views, findings in this study reveal that supervisors are not very narrow-minded, and some of them can appreciate employees' honesty and straight-forward arguments. Therefore, we suggest that employees relieve the prejudice against their supervisors and the concerns over the risks of direct opinion expression. Cynical employees are those who get the worst performance evaluations. Although these employees try to conceal their opinions and avoid communicating problems or dissatisfaction with their supervisors, supervisors can still sense these employees' negative feelings. Since employees' cynicism may result from the long term mistrust and alienation, we suggest that the leadership problems or personality problems behind cynicism should be recognized and addressed.

REFERENCES

References available upon request.

Table 1 Means, standard deviations, correlations, and reliabilities of study variables

Variables	Mean	s.d.	1	2	3	4	5	6	7
1. Voice (self-rating)	5.39	.79	(.90)						
2. Dissent (self-rating)	3.19	1.03	.33**	(.88)					
3. Cynicism (self-rating)	3.42	.78	-.20**	.18**	(.62)				
4. Voice (other-rating)	6.01	.84	.26**	.08	-.16**	(.92)			
5. Dissent (other-rating)	2.88	1.03	-.01	.26**	.14*	.10	(.84)		
6. Cynicism (other-rating)	3.03	.98	-.08	.11	.27**	-.42**	.32**	(.73)	
7. Job performance rating (other-rating)	6.00	1.03	.23**	.08	-.10	.76**	.01	-.47**	(.95)

[a] n = 393. Reliability coefficients for the scales are in parentheses along the diagonal.

[b] Self-rating data were provided by subordinates; other-rating data were provided by supervisors.

* $p < 0.05$

** $p < 0.01$

THE EFFECTS OF MUSIC AND QUEUING INFORMATION ON PERCEPTION OF DOWNLOAD WAITING TIME

Chih-Hui Shieh, National Chiayi University, (886) 273-2824, ch_shieh@hotmail.com
I-Ling Ling, National Chiayi University, (886) 5273-2843, yiling@mail.ncyu.edu.tw

ABSTRACT

Perceived download waiting time is a serious concern. This article is intended to investigate the moderated effects of background music and queuing information on perception of download waiting time. A 3(download delay) × 3(queuing information) ×2(background music) quasi-experiment was conducted with 410 participants. The subjects were exposed to various download delays and webpage designs (queuing information and background music) to examine their perception of waiting time and affective responses. The results showed that online user's perceived waiting time is longer than the actual elapsed time. The longer the delay, the greater was the likelihood to overestimate the waiting period, and also the more biased was their time perception. The findings indicated that providing queuing information will not influence the perception of waiting time. The results also demonstrated that the duration of download delay and background music interacts on perception of waiting time and subjects' affective responses.

INTRODUCTION

Waiting is common in our everyday lives: for example, waiting to get on the bus, to use the ATM, to register a package at the post office or even to use the washroom. In other words, consumer waiting is inevitable in many situations. In early years, scholars and practitioners developed mathematical models to increase efficacy or reduce actual waiting time. However, these models seemed unable to appease impatient clients, especially in the service industries. Arousal of negative emotion is the major contributor to poor service ratings. Maister showed that perceived waiting time becomes a critical factor in consumer satisfaction [19].

Since waiting is unavoidable, reduction in actual waiting time may not be the most compelling issue. Krishna suggests that instead of resorting to scientific measurements, human beings primarily base their perceptions on estimates of size, distance or quantity and thus, a perception bias arises as a gap between reality and perception [17]. Management might find it more efficient in certain circumstances to reap cost-saving benefits by turning this bias to their advantage rather than by improving the overall business operation. Many factors in the environment, such as temperature, colors, music, facilities and information are reported to influence consumers' emotions and time perception when they are waiting [2]. Besides, there is research on waiting at specific locales, such as in line-ups at banks [7], before the check-in desk or at the boarding gate [9].

Advances in Internet technologies over recent years have brought changes to both consumers and vendors by saving consumers a great deal of waiting time and decreasing the venders' costs of running a brick-and-mortar business. Given the exponential growth of revenue in e-commerce, the Internet is a major distribution channel for retailers. As consumers continue to migrate to the virtual world, waiting in the physical environment has shifted over as well. Online waiting as a result of webpage download delay is reported as one of the major annoyances for Internet users [8].

Therefore, this article draws on well-established waiting theories derived in the physical context to probe into consumer affective responses in the Internet context. The objective of this article is to explore the effects of download delay and background music on consumer perception of waiting time and affective responses.

LITERATURE REVIEW

Perception of Waiting Time and Affective Responses in General

Hornik proposes two kinds of consumer time measurement: objective time and subjective time [12, 13]. Objective time is told by clocks or other timing devices in hours, minutes and seconds. Subjective time is subject to the consumers' sheer judgment. Baker and Cameron believe that time perception is how long an individual thinks he has waited, and that one is not aware of being in a waiting state until the elapsed time becomes noticeable [1]. Therefore, keeping consumers distracted may help to prevent them from noticing that time is ticking away. The longer consumers perceive the waiting process to be, the less satisfied they are with the service [15]. Consumers tend to overstate how long they have waited– that is, the waiting period appears longer than in reality [13, 15]. Baker and Cameron emphasize that minimizing the perceived waiting time is critical in evaluating service performance [1].

Most people do not like to wait because waiting brings about displeasure, boredom, anger, anxiety and depression [6]. According to affective responses over the waiting period, Larson concludes that waiting to be served is a typical negative experience which leads to disappointment, depression and worry [18]. Hornik found that positive affect resulted in a different perception of time than negative affect [12]. Consumers underestimated the passage of time when enjoying a good time. They perceived the duration to be longer when they could not wait for something to finish and became anxious about time. Affect dominates consumer behavior. Service failure usually arises when a consumer receives service with negative feelings about the provider [4]. Consumers are likely to walk out if they feel annoyed, edgy and impatient with the waiting process. Consumer behavior is subject to the direct or indirect influence of consumer affect, with positive attitude arising from positive affect and vice versa [10].

Online Waiting

Download time is a key determinant in online waiting and is also considered the most important. It refers to the time period during which an online user waits for a page to load and become usable [23]. Rose and Straub identify six hindrances to the prosperity of e-commerce: download time, measurement of web application success, security weaknesses, lack of Internet standards, limitations in the interface, and requests for hypermedia [21]. Guynes asserts that excessive download time brings about negative responses in users who have urgent tasks to perform [11]. It is even more frustrating when a website cannot accommodate anymore visitors or a webpage simply does not open [3]. Wirtz and Bateson showed that delays in the system result in criticisms and thus poor satisfaction reviews of the system [25].

Over the years information systems have greatly improved in response time to reduce download delays. There are still many other drags on the loading speed, such as insufficient bandwidth on either end of the Internet connection, oversize files, and inadequate hardware or software. Unfortunately, as much as service providers would like to do everything they can to reduce download time. They can only upgrade the server and keep the items on the webpage reasonably-sized. It is beyond their power to manage the end user's equipment and delivery of dataflow [20]. System operators often minimize web pages or upgrade hardware to solve severe lags. Sometimes these efforts are unaffordable or their realized benefits are negligible relative to the investments [22].

Fill Mechanism

Fill mechanism distracts consumers and interferes with their perception of time [19]. The aforementioned literature on time perception posits that people feel they have waited for a long time if they have been mindful of time, but distraction may reduce their perception of the length of time. The purpose of fill mechanism is to engage waiting consumers physically and mentally in order to divert their awareness of time. With appropriate fill mechanism in place, the consumers' evaluation of service is more likely to be favorable [24].

Kellaris and Kent believe that stimuli such as music in the service surroundings can impact consumers' emotions and alter their mental state [16]. Baker et al. also consider music to be a stimulus in a storefront and, along with other stimuli, to have a significant effect on consumer perception and repurchase intention [2]. Studies across a wide array of disciplines have documented music's effect on listeners' emotions and thoughts. In a real business environment, consumers' positive or negative affect caused by music impacts their temporal perceptions [5].

Based on the resource allocation model, some scholars believe that the provision of delay information may reduce both the consumers' need to keep track of time and their perception of waiting duration [15, 24]. The consumer's satisfaction level drops when there is a continuous provision of delay information such as how long one has to wait or how many people are waiting [18, 19]. Hui and Tse categorize delay information into two kinds: delay information which is *waiting-duration information* and *queuing information* [14]**.** The former means an estimate of the time span before one is served or how much longer a consumer has to wait; the latter represents the sequential place of a consumer on the waiting list or how many people are waiting. Their research also argues that the effect of delay information varies depending on the length of the lineup. Therefore, providing information is not necessarily the most effective means to alleviate consumer dissatisfaction.

In general, waiting in line for a service undoubtedly frustrates, depresses and antagonizes consumers very easily [18]. Consumers make social comparisons when they are in a line-up. Namely, their mood is affected by the number of people waiting ahead of and behind them. The fewer ahead and the more behind, the less likely they are to quit waiting and positive affect increases as well [26]. The purpose of this study was to find out whether different levels of download delay (short/medium/long) and playing background music make any difference to user perception of waiting time and affective response.

METHOD

Participants and Design

A sample of 410 participants was generated via Internet websites. It employed a 3×2×2(download quickness× queuing information× background music) between-subjects design. The download quickness was manipulated at three levels(5-second, 30-second, or 60-second). The queuing information was manipulated at three levels(0 ahead, 5-people ahead or 30-people ahead). The authors shuffled the screen webpage to randomly mix the eight controlled conditions before distributing them to participants.

Procedures

A popular movie was selected as a stimulus to encourage online users to explore a specifically designed hyperlink. A hyperlink was posted in a movie share forum, a click on which allegedly enabled net users to download a Flash animation of the movie, *Pirates of the Caribbean: at World's End*.[1] The

---

[1] Based on protecting intellectual property rights , our experiment was a fictitious scenario to attract the online users to visit the designed hyperlink.

participants began the online waiting experiment by clicking to download. The movie Web pages were presented to the users with different conditions depending on the purpose of the investigation.

Manipulation Check and Measures

A 0-second quickness is unrealistic for a page to load as it takes even the fastest Internet connection a couple of seconds. Based on Rose and Straub (2001) and Rose et al. (2003), delay levels of 5 seconds(control), 30 seconds, and 60 seconds were used as the treatment level. Differences in delay attitude indicated which delays would be used for the two treatments along with the control of 5-second delays.

A manipulation check on attitude toward the delay was conducted. The subjects were asked to rate the perceived feeling for download quickness on a 5-point scale ("excessively slow," "slow," "acceptable," "fast," "excessively fast,"). Perceived wait time was based upon each participant's estimate of how many seconds it took for the page to download completely.

For the feelings of negative affect measure that follow, the participants were asked how the felt at the time they were waiting for the film to download. The subjects were asked to indicate how well each of six statements described how they felt in this circumstance (1="not at all" ; 5="very much so"). The statements were: I feel calm, annoyed, happy, relieved, impatient, and anxious. Three positive items(calm, happy, and relieved) were reverse-scored to construct the dependent measure. The coefficient alpha for this measure (α =.80) represents a sufficiently high level of internal consistency. At the end of the experiment, the subjects answered questions regarding their age, gender, and Internet experience.

RESULTS

Manipulation Checks

It was found that a significant difference existed for the subjects among these three levels of download quickness ($M_{5\text{-}sec}$=3.31, $M_{30\text{-}sec}$=2.20, and $M_{60\text{-}sec}$=1.85; $F(2,407)=87.33$，$p<0.001$). Post hoc tests showed that all means were different from each other at the 0.05 level. The results indicated that the manipulation was successful.

Effects of Fill Mechanism on Perceived Waiting Time and Negative Affect

The means values for perceived waiting time and negative affect in different situations were shown as in Table 1. We performed a 3(DQ(download quickness):5-sec, 30-sec, and 60-sec) × 3(QI(queuing information:0-people, 5-people ahead, and 30-people ahead) × 2(BM(background music: with vs. without))between-subjects ANOVA. In terms of the effect on perceived waiting time, the results reveal a main effect of DQ($F(1,409) = 74.24$, $p < .001$) and background music ($F(1,409) = 6.35$, $p < .05$). There is no significant main effect of queuing information($F(2,409) = 0.26$, $p > .1$) on perceived waiting time. Not surprisingly, perceived waiting time is higher when the download quickness is 60-sec (long) than when it is 30-sec (medium) and 5-sec (short). Participants who were given a background music condition perceived longer perceived waiting time than those who were given a mute background.

Furthermore, the DQ × BM interaction reached significance ($F(2,392) = 3.02$, $p < .05$). As predicted, the background music(BM) effect emerged as significant for 30-sec and 60-sec conditions but not for 5-sec condition. In addition, the QI × BM interaction reached slight significance ($F(2,392) = 2.86$, $p < .1$) As predicted, the background music(BM) effect emerged as significant for 5-people ahead but not for 30- people ahead condition. The consumers in the background music condition perceived a longer waiting time (second) in the 5-people ahead condition (M = 92.96) compared to in the no-wait condition(M=83.65) and in the 30-people ahead condition (M =73.31). For mute condition, there was no significant difference between the no-wait condition and 30-people ahead condition (M = 71.25 vs. 74.18).

In terms of the effect of negative affect, there is a main effect of DQ($F(2, 409) = 9.46$, $p < .001$), such that the longer download quickness, the lower the negative affect($M_{60\text{-sec}} = 3.11$ vs. $M_{30\text{-sec}} = 3.04$ vs. $5_{5\text{-sec}} = 2.62$). However, an ANOVA with negative affect revealed a non-significant main effect of queuing information($F(2,409) = 0.54$, $p > .1$) and background music(($F(1,409) = 0.93$, $p > .1$) Meanwhile, results indicated a slight tendency for a two-way interaction ($F(2,392)=2.57$, $p<.1$).When the situation of download delay was 5 seconds and 30 seconds(low and medium levels), the subjects had less negative affects when the background music was played. The reverse tendency was observed for the high level of download delay ($M_{\text{music}}=3.00$ vs. $M_{\text{mute}}=3.21$).

DISCUSSION

Across our study, it was found that download quickness is a critical trigger for user perception of waiting time and affects. As the quickness duration extends, the user's negative affects become stronger. In the meantime, the difference between actual and perceived time widens, and so temporal perception bias becomes more serious. Even though quickness information reduces uncertainty, this effect is offset by increasing impatience as the download duration extends. As a result, negative affects begin to take hold.

Additionally, download quicknesss were experimented with in conjunction with background music. It was found that only at a comfortable level of quickness (for example, under 30 seconds) were the effects of music significant on user affects. As the quickness extended up to 60 seconds, users became tired of the music and their negative emotions increase. When presented with a long quickness (60 seconds) with music, the average of perceived time nearly increased 75 seconds compared to the same time period without music. Therefore, when download quickness is too long, fill mechanism has no effect on user perception of time and positive affects; instead, it may even incur a backfire effect.

This study has observed that download quickness remains critical for online user perception of waiting time and affects. At all three different levels of download quickness, perceived waiting time on average is longer than the actual duration, which indicates that consumers are very likely to have a temporal perception bias while they are waiting online. They tend to overstate the time duration of waiting for service. Besides, such bias becomes more significant as the quickness extends.

Users have the tendency to perceive the duration to be longer than it actually is. As the quickness extends, such perception bias becomes more serious. This finding implies that consumers might be more impatient online when they have contact with only the computer image than they are on the premises. Temporal perception bias is more likely to arise, and so are negative affects. Online service providers ought to continue with their efforts to improve loading speed and fill mechanism. In addition, previous studies on waiting suggest that fill mechanism influences consumer perception of waiting time and affects on the premises. However, our findings have found that in an online environment, fill mechanism is effective to various degrees depending on the duration of download quickness. Playing music on the webpage has no effects on user perception of waiting time. Besides, even though music influences user affects, its effect is significant only at the medium level of download quickness. Our findings suggest the need for the system manager to keep download quickness at a reasonable level. Otherwise, providing music will not make a difference.

The mechanism of background music is obvious at the comfortable level of quickness for 30 seconds. When the quickness is as long as 60 seconds, consumers become used to it and are rather annoyed by it. The use of music may backfire at the vendor because it exacerbates the user's anxiety or impatience. Therefore, in an online environment, download quickness remains critical for consumer perception of waiting time, affects and willingness to wait. Fill mechanism is most effective at a reasonable level of quickness. When the duration is too long, this mechanism might create undesirable effect

This study was intended to investigate truths of online waiting, and therefore our experiment was performed under authentic circumstances. This study was unable to ensure that the subjects were not distracted by uncontrollable factors while waiting for the page to load. Additionally, our study only investigated the two factors of download quickness and background music. Future research can explore other quickness information such as whether the number of users waiting ahead or behind will influence users' willingness to wait or not.. Also, other kinds of fill mechanism that future research can investigate are web banners and information of download speed on user perception, affect and willingness to wait.

REFERENCES

[1] Baker, J., and Cameron, M. The effects of the service environment on affect and consumer perception of waiting time: an integrative review and research propositions. *The Academy of Marketing Science*, *24,* 4 (1996), 338-349.

[2] Baker, J., Parasuraman A., Grewal, D.; and Voss, G.B. The influence of multiple store environment cues on perceived merchandise value and patronage intentions. *Journal of Marketing, 66*, 2 (2002), 120-141.

[3] Barber, RE., Lucas, H.C.J; and Kling, R. System response time operator productivity, and job satisfaction. *Communications of the ACM*, *26*, 11 (1983), 972-986.

[4] Bitner, M.J. Evaluating service encounters: The effects of physical surrounding and employee responses. *Journal of Marketing, 54*, 2 (1990), 69-82.

[5] Bruner, G.C. Music, mood and marketing. *Journal of Marketing, 54*, 4 (1990), 94-104.

[6] Carmon, Z., Shanthikumar, J.G.; and Carmon, T.F. A psychological perspective on service segmentation models: The significance of accounting for consumers' perceptions of waiting and service. *Management Science.41,* 11 (1995), 1806-1816.

[7] Chebat, J.C., and Filiatrault, P. The impact of waiting in line on consumers, *International Journal of Bank Marketing, 11*, 2 (1993), 35-40.

[8] Ewalt, D.M. Count to 10, then read this. *InformationWeek*, http://www.informationweek.com/story/WK20020219S0013, Accessed on 2009/05/19

[9] Folkes, V.S., Koletsky,S.; and Grahma, J.L. A field study of causal inferences and consumer reaction: the view from the airport. *Journal of Consumer Research, 13*, 4 (1987), 534-539.

[10] Gardner, M.P. Mood states and consumer behavior：A critical review. *Journal of Consumer Research, 12*, 3 (1985), 281-300.

[11] Guynes, J.L. Impact of system response time on state anxiety. *Communications of the ACM, 31* , 3 (1988), 342-347.

[12] Hornik, J. Situational effects on the consumption of time. *Journal of Marketing, 46*, 4 (1982): 44-55.

[13] Hornik, J. Subjective vs. objective time measures: A note on the perception of time in consumer behavior. *Journal of Consumer Research, 11,* 1(1984), 615-618.

[14] Hui, M.K., and Tse, D.K. What to tell consumers in waits of different lengths: An integrative model of service evaluation. *Journal of Marketing, 60*, 2 (1996), 81-90.

[15] Katz, K.L., Larson, B.M.; and Larson, R.C. Prescription for the waiting-in-line blues: entertain, enlighten, and engage. *Sloan Management Review, 32,* 2(1991), 44-53.

References available upon request.

DOES THE PROSPECT THEORY NECESSARILY MEAN DISPOSITION EFFECT? AN EXTENSION OF DISPOSITION EFFECT

Kuo, Min-Hua, Shih Hsin University (886-2-) 2236-8225 ext 63435, mhkuo@cc.shu.edu.tw

ABSTRACT

Derived from the prospect theory, the disposition effect argument argues that investors ride losers too long and sell winners too soon. A lot of evidence has lent support on them. However, both of the theory and the argument leave questions to be solved. First, why do the investors in real markets sell losers then? And when do they sell? Second, why do the investors in real markets not sale winners immediately when the stocks become winners? And for how long do investors hold winners before selling them? We answer these questions by applying the theory of mental account, and show that the disposition effect is just one of different disposition patterns complying with the prospect theory.
Keywords: maximum loss tolerance, minimum value threshold, disposition effect, prospect theory, mental account

INTRODUCTION

Kahneman and Tversky's (1979) prospect theory is one of the most important behavioral models regarding decision making under risk. According to the theory, the value function is convex in the loss region, and concave in the gain region. That is, agents are risk seeking when facing losses and risk averse when facing gains. Shefrin and Statman (1985) apply it to investment behaviors and argue that investors are thus more disposed to sell winners instead of losers. They name it as the disposition effect. Since then, abundant empirical evidences have lent great supports to it. However, both the prospect theory and the disposition effect argument leave puzzles to be solved. The convexity implies that, as long as the prices of losing stocks are probable of rising again, investors will never realize the loss, given other conditions being constant. The concave implies investors would realize gains immediately when asset prices go up over the cost. These arguments are certainly inconsistent with what we see in real markets. A gap between theory and practice needs bridging. The unsolved questions to be addressed in this study are as follows. First, when will investors realize losses? Or, borrowing from the phrase of disposition effect, how long will investors ride losers? Second, how early will they sell winners? Appling the theory of mental account to answer these questions, we propose people use different decision principles for mental accounts of losses and gains respectively. With this structure, we answer the questions mentioned without changing the value function. The results indicate several disposition patterns complying with the prospect theory; the disposition effect is just one of them. The next section will briefly introduce the prospect theory, disposition effect and the mental account theory, followed by the discussion of our main arguments. We use a hypothetic case to illustrate the idea. The conclusion is in the last section.

LITERATURE REVIEW

Kahneman and Tversky (1979) assume that value function V drives an agent's decision:

$V(x, p) = \Sigma v(x)w(p)$

in which v(x) is the subjective value of the outcome x, and w, a function of x's probability p, represents the "attractiveness" of outcome x to the agent. The value function v(◦) is concave over gains and convex over losses. They lead the agent to be willing to bear risk when a loss happens, in the hope of gaining back, and unwilling to bear risk when a gain shows, in the fear of gain vanishing. Kahneman and Tversky use experiments to demonstrate that subjects are prone to take a risk over losses in order to avoid loss realizations, while prefer certainty over gains. Shefrin and Statman (1985) apply the prospect theory to investment and argue that investors tend to sell winners too early and ride losers too long. To test it empirically, Shefrin and Statman (1985) use individual investor's stock trading and aggregate mutual funds data to measure the tendency of disposition effect for the market as a whole. The measures, constructed as the number of transactions with gains over that of all transactions, are reported around 0.60 whatever the round-trip durations are. That is, winner stocks are more probable to be sold than losers, in consistent with the prospect theory.

Since then, the evidences in support of the disposition effect have been piled up rapidly (such as Odean, 1998; Shapira and Venezia, 2001; Weber and Camerer, 1998; Grinblatt and Keloharju, 2000; Shapira and Venezia, 2001; Shu, et. al, 2005). The impression seems to be that the disposition effect dominates investors' behavior in financial markets. For example, Odean (1998) uses 10,000 customer accounts randomly selected from a nationwide security house in the U.S. to test whether investors sell winners too early and hold losers too long. His findings lend strong supports to the disposition effect because the proportions of stocks having gains realized during a yearly (PGR), 0.148, is significantly greater than that of loss realized (PLR), 0.098. It indicate a winner stock being at least 50 percent more likely to be sold from day to day than a loser stock, which is similar with Weber and Camerer's (1995) experimental results. In sum, investors demonstrate a strong preference for realizing winners rather than losers.

The prospect theory and disposition effect are very valuable in describing investment behavior. However, some issues remain further insightful thinking. Given the assumption of convexity and concavity, individuals have no motivation to realize losses and to keep winners for some time. It is certainly not the case in real markets. There are 40% realizations are loss realizations in Shefrin and Statman's empirical reports; about one-tenth of losers in average are realized in Odean's reports; the average yearly return on realized gains is 27.7%. Questions: When do investors dispose their losing stocks? For how long do investors hold winners before sell them? Do these "anomalies" mean the value function not working? We apply metal account theory to explain these questions.

PROPOSITIONS

According to the prospect theory, people behave different in loss regions from gain regions. It is reasonable thus to infer that people are used to set up accounts for gains and losses respectively, and adopt different decision criteria in two types of accounts. We propose that for the decisions of hold-or-sell, investors use maximum loss tolerance as the decision criterion in loss accounts and use minimum value threshold in gain accounts. Investors do not realize losses until the loss reaches maximum loss tolerance of them; they do not sell winners unless the (subjective) value from gain realization overpasses the minimum value threshold. The basic assumptions are that people do not tolerate "too-big loss", the loss beyond their limits; they would assume the risk of holding losers only when the potential loss is below the loss limit. Likewise, people do not feel satisfied with "too-little gain", the gain below the threshold; they would sell it only when the gain realized is greater than the threshold.

The loss tolerance is a financial mechanism, while the value threshold is a psychology mechanism. People are subject to both financial and emotional (or objective and subjective) conditions when making decisions under uncertainty. However, only subjective risk preference is considered in the structure of prospect theory. Financial resources are limited, so the loss one can allow is limited up to a certain level. Every investor has a maximum loss tolerated, explicitly or implicitly. Likewise, a minimum value must be generated from realizing winners, otherwise the transaction will not be intrigued. Therefore, we suggest that the decision of hold-or-sell is up to one's maximum loss toleration in loss mental account and minimum value threshold in gain mental account. We will illustrate this by a hypothetical case and figure 1.

Assumed investor A's maximum loss allowance in the loss mental account of is $-\alpha$, whose corresponding value is $-V_{\alpha}$. When the price drops, he/she will not sell the losers, due to the convexity of value function, until the payoff reaches $-\alpha$. Any losses greater than $-\alpha$in absolute terms) is overpass his/her financial capacity. So the maximum drop investor A allows is $-\alpha$. On the other hand, in the gaining mental account, the minimum satisfaction (value) required is $V\alpha$, whose correspondent payoff is α. That is, any satisfaction below V_{α} is not great enough for investor A to sell winners, until payoff greater than α. Our extension by mental account theory requires no assumption change regarding value function. People complying with the value function as proposed in prospect theory will dispose losers when the payoff touches the maximum loss tolerance. Similarly, they will wait before the winners pass the minimum value threshold. In other words, investors' maximum loss tolerance determines how long they will keep the losers; and minimum gain threshold determines how soon to sell the winners, with other things being constant.

By this structure, we get a few insights. Since the maximized loss and the minimum value threshold are different across investors, investors demonstrate different disposition patterns. The disposition effect is only one of them. In the case of investor A, the size of $-\alpha$ is greater than α, he is more likely to demonstrate a disposition effect, riding loss very long while selling winners very soon. However, in the case of investor B, the size of $-\beta$ is smaller than β, he is more likely to demonstrate a "reverse" disposition effect, riding losses shortly while selling winners late. If the sizes over loss and gain are symmetrical, then a third pattern exists: symmetrical disposition pattern. For those without emergent financial constraint, maybe due to the small investment size, they are not forced to realize losses and would prefer not realizing any losses, which is consistent with the prospect theory.

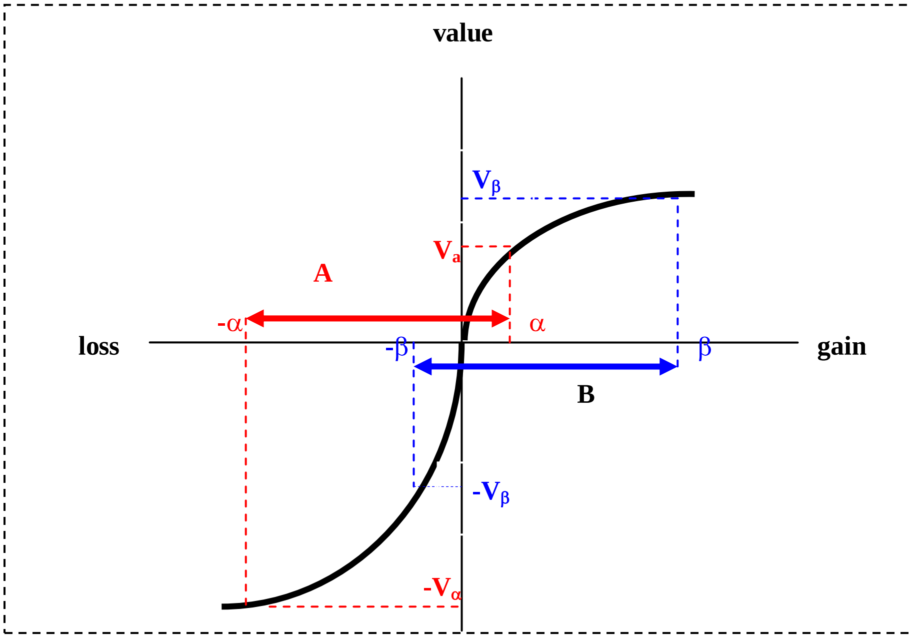


FIGURE 1

The strength of our proposition is bridging the gap between the prospect theory and the real disposition pattern we observe in the real market, and not requiring any change in the value function. If an investor complies with the value function in the prospect theory, he/she is still likely to demonstrate a contrasting behavior pattern against the disposition effect: he/she may sell losers quickly and keep winners much longer, which we call the "reverse disposition effect", or a symmetrical disposition pattern. Take two hypothetical investors A and B in figure 1 as examples. Both of them are assumed to have the same value function, concave for gains and convex for losses, and every relevant feature is identical between them except the maximum loss tolerance and minimum gain threshold. For example, investor A can tolerate more pain (and loss) than B, and require more gain to be satisfied subjectively. That is, $-V_{\alpha}<-V_{\beta}$ and $-\alpha<-\beta$; $V_{\alpha}<V_{\beta}$ and $\alpha<\beta$. That is, investor A allows himself to bear uncertainty until the loss reaches – , but investor B only to the extent of $-\beta$. A's maximum loss tolerance is greater (in absolute terms) than B's. On the other hand, investor B requires a higher profit than A in realizing the gains before he/she feels satisfied with the transactions; $\beta>\alpha$. B's minimum gains threshold is higher than A's. Assume each buys a certain stock at $50. How long will they hold onto it before selling? According to the disposition effect, it depends on whether the stock is in gain or in loss. They will hold longer if the stock devalues rather than increases, because they are risk-seeking over loss area and risk-averse over gains, as long as the investors behave in accordance with the Prospect Theory. They will not get rid of the loser until its value decreases up to, say, 15%, or until it gains up to 10%. Investor A and B will demonstrate the same pattern: riding losers long and selling winners soon. However, we consider this perspective incomplete in its lack of accommodating the differences in loss tolerance and gain threshold. According to figure 1, how long the investor will ride a loser depends on the maximum loss tolerance or the maximized loss he/she can accept; and how quickly he/she will sell a winner depends on the minimum gain required, which is the gain threshold to sell. If the investor will not tolerate any loss greater than 5%, such as investor B, he/she will sell it before the loss goes over the limit. And if B is not satisfied with a gain less than 15%, he/she will not sell it before this level. In this case, he/she exhibits a shorter length in holding a loser than holding a winner, even if the investor's value function is consistent with the Prospect Theory.

Based on the reasoning above, our first main proposition is that both disposition effect and Reverse disposition effect coexist in financial markets, even if the Prospect Theory applies.

Table 1 A Hypothetical Case Illustrating the Potential Overestimation of the disposition effect

Five hypothetical investors are assumed to own two \$5-losers and two \$5-winners respectively. But the realization decisions are different. A sold all losers and one winner; B sold one loser and all winners; C sold one winner and one loser; D sold none; and E sold two winner. D and E are Strict Loss-Aversion and refuse to realize any losses. All numbers are in dollars. SS's ratio represent the measure developed by Shefrin and Statman (1985), with the transaction number of gains realizations over all transactions. Odean's ratios, the PGR and PLR, are the proportions of gains (losses) realized, which is the ratios of realized gains (losses) over the sum of realized gains (losses) and paper gains (losses).

INVESTORS	A	B	C	D	E	SS'S RATIO	ODEAN'S RATIOS
POTENTIAL LOSSES (2 STOCKS)	10	10	10	10	10	6/10 =0.6	PLR= 20/50=0.4
REALIZED LOSSES	10	5	5	0	0		
POTENTIAL GAINS (2 STOCKS)	10	10	10	10	10		PGR= 30/50=0.6
REALIZED GAINS	5	10	5	0	10		

Based on the discussion above, the market seems having a group of investors with the attitude of strict loss aversion, and deserves more profound examination. Although it is possible that such attitude is just what the disposition effect describes, we need to make empirical scientific analysis before supposing or opposing it, rather than ignoring it and simply including it in the pattern of the disposition effect.

Let us make the discussion with figure 1. For any investor i,

if ai > αi, the investor tends to be of the disposition effect;

if ai < αi, the investor tends to reveal the Reverse disposition effect;

if ai = αi,, the investor tends to be rational in terms of the symmetry over gains and losses.

In addition, if loss is not tolerable at all, the mental account of loss is closed and the investor demonstrates the fourth type of pattern of strict loss aversion. In other words, the attitude of refusing to sell losers is not because such investors can tolerate greater loss; quite to the contrary, they cannot tolerate failure.

We will empirically examine, on the level of individual investors, to get a picture of disposition patterns in practice. Aggregation analysis allows us to know the investors' disposition as a whole, but at the expense of knowing more of different investment behavior patterns. We find it urgent to know whether or not the disposition effect really dominates the investors. How many investors are disposed to selling winners? And how many refuse to realize any losses? How many investors are rational, in terms of their disposition of selling stocks to be symmetrical over gains and loss areas? In the current behavioral finance mainstream, the disposition to sell winners seems to be dominant. However, from the demonstration in the section of Figure 1, it is highly possible that the investors demonstrate a reverse disposition effect, even when they comply with the value function of the Prospect Theory. Not only are these questions interesting, they are very valuable in developing alternative positive theories to describe how investors actually behave under uncertainty.

Since investment is a behavior under a veil of uncertainty, what people do NOT do is as important as what people do. For example, selling a certain stock will add pressure on its price; NOT selling it will lessen the pressure in the bear market. Both are equally important in determining the stock price. To tap into market behaviors more accurately, we need an approach to explore what investors are really thinking, instead of what they already have done. Questionnaires and laboratory experiments, borrowed from psychology, are two approaches that have earned increasing attention in the finance field ever since the emergence of behavioral finance. Due to the limits of sample scales, laboratory experiments in many issues are considered less representative than the questionnaire survey on the larger scale. This study utilizes a large-scale survey to explore uncertain investment behaviors over gains and losses.

CONCLUSIONS

The disposition patterns have been an attention-catching issue for more than two decades, since Shefrin and Statman extended Kahneman and Tversky's prospect theory to their analysis of investment behaviors. This is widely supported by empirical results. However, we need to address the questions unanswered: consider the maximized loss an individual investor is willing to tolerate and the minimized gains required for her/him to feel satisfied. When this concept is taken into account, we find that both disposition effect and Reverse disposition effect are possible, given the value function suggested by Prospect Theory. What is worth to discuss here is that the mechanisms behind these two constructs are different. The mechanism of maximum loss tolerance is economic factor, while that of minimum gain threshold is psychology factor. People making decisions are subject to financial and emotion constraints, or objective and subjective constraints. When dealing losses, people are tend to avoid losses but subject to financial conditions. Without financial consideration, it is possible that people just shy away from realizing losses. In the similar way, how far is the gain required above reference point (which could be the purchase price or the cost of capital or others, depending on the individual's preference.) is basically a matter of psychology. People need to subjectively feel satisfied, getting the normal return is not enough for many people. That is why investors do not sell winning stocks until it reach some return threshold.

Secondly, we argue that the measurement prevalently used ignores an important group of investors refusing to sell losers and includes it in that of disposition effect, which has been demonstrated inadequate by the evidence that is presented in this research. In sum, by incorporating the loss tolerance and the gain threshold, we offer a new perspective on considering disposition patterns and enrich the behavior patterns implied by the Prospect Theory. At least four disposition patterns are demonstrated by our empirical analysis.

REFERENCES

References available upon request from the author.

THE EFFECTS OF ONLINE FEEDBACK RATINGS ON CONSUMERS' TRUST AND RISK PERCEPTION

Yung-Chuan Lin, National Chiayi University, (886) 273-2824, chyuan17@gmail.com
Chung-Chi Shen, National Chiayi University, (886)5273-2910, ccshen@mail.ncyu.edu.tw
I-Ling Ling, National Chiayi University, (886) 5273-2843, yiling@mail.ncyu.edu.tw

ABSTRACT

This article focuses on how feedback sources influence consumers' risk perception and trust. A 2(rating ratio) × 2(negative rating density) ×2(rating framing) online experiment was conducted with 214 participants. The results indicate that when consumers consult the feedback sources, they are more influenced by the negative rating density than the rating ratio source. More importantly, the type of risky choice framing also had a significant influence on the judgment of trust and risk perception. Combining the rating framing effect, the three feedback sources create an interaction effect. Theoretical and managerial implications of the findings are provided.

INTRODUCTION

Traditional research on the influence of recommendations on consumers has typically been subsumed under personal influence or word-of-mouth (WOM). Consumer reviews based on opinions and shared experiences in the use of a product is a powerful source of information about consumer preferences that can be used in recommender systems. Additionally, opinion leadership and reference groups also relate to the study of recommendations and to influence in general [1]. The online recommender system was more influential than traditional recommendation sources such as human experts and other consumers [2].

Online auction is an electronic commerce application. Traders can sell new or used merchandise via this platform to facilitate resource redistribution. Compared to offline shops, online auctions significantly slash costs, such as rents, human resources and logistics. Many startups are built on this model. Some offline shops have also started to participate in online auctions in order to increase their consumer base. As the online auction market gradually matures, auction sites develop new systems, such as "Buy Now" in Yahoo!'s Auctions, skipping the bidding process, to enhance efficiency. The online auction has become an e-commerce platform aggregating diverse product information.

Although the online auction has numerous advantages over traditional auctions, it creates far more risk perception than offline trade due to the virtual environment [3]. Consumers' risk perception becomes the major obstacle to e-commerce development[4]. Therefore, many auction sites establish feedback systems between consumers and sellers as the foundation for mutual trust[5]. Rai, Patnayakuni and Seth (2006) indicate that a feedback system can facilitate problem-solving and communication, thus improving overall quality and performance.

The ability to exchange opinions and experiences online is known as online word of mouth (eWOM) and has been shown in the literature to have the potential to impact e-commerce sales[6]. Online auctions have expanded quickly in recent years. Facing thousands of online sellers, consumers often establish their initial understanding and trust via a feedback system. The main objective of this study is to investigate the influences of online feedback sources on consumers' online trust evaluation and risk perception. In addition, we explore the moderating effect of the framing message related to feedback sources and online trust and risk perception.

LITERATURE REVIEW

Understand the Influence of Feedback System on Online Transactions

A feedback system is developed by the auction sites to evaluate sellers' performances. That these records are public and cannot be deleted unilaterally becomes a critical element to transaction decisions. Sellers and consumers can provide feedback and recommendations based on their transaction experiences. The system records every transaction. Therefore, ratings accumulated can be a reference for credibility. A feedback system is an incentive for sellers to maintain positive credibility [7]. The content of the recommendations influences sellers' credibility and positive comments can be assessed as a determinant to online transactions[8,9]. Anonymity and insufficient technology applications may undermine mutual trust and confidence during transactions[10]. Even though rating is an important reference for consumers to evaluate sellers' credibility, consumers still do business with sellers within a certain ratio of negative feedback[11].

Online recommendation sources range from traditional sources such as other consumers (e.g., comments of consumers on retail websites) to personalized recommendations provided by the feedback rating systems (e.g., eBay's DSR system). The number of feedback responses significantly affects consumers'' image toward sellers and final price[5,10,12,13]. Consumers could use the following information sources for their next durable goods purchase: (1) positive comments[8], (2) negative feedback rating[14], (3) the number of feedback responses[5].

The current study focuses on two determinants that could influence the impact of computer-mediated feedback on consumers' online evaluations: positive rating ratio and negative rating density. We posit that these online feedback sources will have different levels of influence on consumers' trust and perceived risk. We aim to analyze whether the positive rating ratio and negative rating density will decide the consumption simultaneously. Prior research has shown that higher positive feedback leads to higher trust towards sellers. On the other hand, under the same feedback system, we try to verify whether a higher negative feedback density will change the consumers' trust towards sellers. When two sellers have comparable positive/negative feedback, one of them will receive less trust from consumers if he/she receives more negative ratings in the recent cycle. The number of feedback responses has a significant influence on the level of trust, which is critical to online transactions.

Based on the preceding review of the literature we postulate that positive rating ratio as well as negative rating density providing product feedback information will influence consumers in computer-mediated environments [11]. We thus formulate the following hypotheses.

H1: For low negative rating density, consumers will perceive a higher trust toward a seller in good rating ratio condition than in bad rating ratio condition.

H2: For low negative rating density, consumers will perceive a lower risk perception toward a seller in good rating ratio condition than in bad rating ratio condition.

Risky Choice Framing Effects

The framing effect refers to people changing their preferences based on two different phrasings of the same question. The framing theory is applied in a wide range of research areas with varied results. When people change their preferences due to different statements (positive/negative) of two identical decisive questions, this is called the “framing effect”[15]. Different framings will influence how receivers understand and evaluate information [16].

“All frames are not created equal.” This was the theme of an article in which it was proposed that there are three types of framing effects that are distinguishable in terms of their operational definitions, their typical results, and the likely underlying process[17]. The article summarized earlier research in which objectively equivalent information resulted in different judgments and decisions depending on the way in which the information was *framed*. The authors identified three distinct types of framing effects in the literature: attribute framing, goal framing, and risky choice framing.

Risky choice framing occurs when willingness to take a risk (e.g., a medical treatment with potential outcomes) depends on whether the potential outcomes are positively framed (e.g., life-saving rate) or negatively framed (e.g., life-losing rate). Based on risky choice framing, this study tries to analyze the influence of information framing towards online feedback rating evaluation. Under positive framing, information will connect to comments that memory supports, whereas information under negative framing will connect to comments that memory opposes. We formulate two hypotheses which consider potential reasons for which various risky choice framing effects may differ in their influence on consumers’ judgments. Thus, the following hypotheses are posited.

H3: For high positive rating condition, consumers will perceive a higher trust toward a seller in positive framing terms than in negative framing terms.

H4: For high positive rating condition, consumers will perceive a higher risk perception toward a seller in positive framing terms than in negative framing terms.

METHOD

Participants and Design

A sample of 214 participants was generated via Internet websites. A 2(rating ratio: good/bad) × 2(negative rating density: high/low) × 2(rating framing: positive/negative) between-subjects online experiment was conducted. The authors shuffled the screen webpage to randomly mix the eight controlled conditions before distributing them to participants.

Procedures and Measures

The experiment begins with a click on the hyperlink posted at the online platform “Youthwant”. The participants are directed to the greeting page (page 1) of *Youthwant Feedback*. On the left is a notice, as on page 2, requesting users to read the online feedback rating system in the seller’s portfolio and to evaluate their trust and perceived risk toward the seller before making their purchase. The experiment begins with a click on “Click to evaluate”.

Manipulation

To check whether the feedback sources were manipulated successfully, two manipulation checks were conducted. The rating ratio refers to the percentage of positive ratings in all feedback. Based on a pre-test, a good ratio is 99.22%, and a bad ratio is 90.4%. The second between-subject factor manipulated the negative rating density. This refers to the density of negative feedback ratings in one month and in six months. It was adopted from two locally renowned online auction websites, Yahoo! and Ruten (i.e., eBay’s). Negative rating was categorized as high and low. High density represents the negative ratings within one month, whereas low density represents the negative rating over the past six months. Finally, the last factor was rating framing. Rating framing refers to the presentation of feedback rating on web pages. It is presented as positive rating or negative rating. Positive framing emphasizes the positive rating ratio (e.g., positive rating ratio: 90.4%), while negative framing emphasizes the negative rating ratio (e.g., negative rating ratio: 9.6%).

After reviewing the experimental information using a 7-point Likert scale ((1=strongly disagree; 7=strongly agree), the participants were asked to answer two questions: "What do you think about this seller's rating ratio?", "What do you think about this seller's negative rating density?"

Measure

To assess participants' risk perception while they were reading the feedback information, the participants were asked to indicate how they felt about each of six statements (1=strongly disagree; 7=strongly agree). Jacoby and Kaplan's (1972) scale was adopted to evaluate risk perception. Factor analysis produced 9 items into one single factor (73.68%), and the reliability is high (α = .95). Furthermore, trust towards the seller was evaluated. The measurement scale was adopted from McKnight and Chervany (2002). The results were evaluated by a 7-point Likert scale. Factor analysis was used to produce one single factor (71.96%) from the 7-item scale, and the reliability is also high (α = .94).

RESULTS

Manipulation Checks

Following the online feedback task, participants involved in the experiment were asked to judge the type of rating ratio and negative rating density before they shopped. It was found that a significant difference existed for the participants among two levels of rating ratio (M_{good} = 6.12, M_{bad}=1.96; $F(1,213) = 1165.13$, $p < .01$). The results indicated that the manipulation of the rating ratio was successful. The results also indicated that participants who were assigned to the high negative rating density condition got higher scores than those who were in the low negative rating density (M_{high} = 5.98, M_{low}=1.40; $F(1,213) = 663.89$, $p < .01$). Thus, the negative rating density manipulation was effective.

Effects of Feedback Sources on Trust and Risk Perception

The means values for trust and risk perception in different situations were shown as in Table 1. We performed a 2(RR(rating ratio): good vs. bad) × 2(NRD(negative rating density: high vs. low) × 2(rating framing: positive vs. negative) between-subjects ANOVA. In terms of the effect on trust, the results reveal a main effect of RR($F(1,206) = 100.47$, $p < .001$), NRD($F(1,206) = 674.70$, $p < .001$), and rating framing ($F(1,206) = 13.10$, $p < .001$). Not surprisingly, trust is higher when the rating ratio is 99.22.% (good) than when it is bad (91.4%). Participants who were given a higher negative rating density condition perceived lower trust than those who were given a lower negative rating density. Positive rating framing is more significant on trust than negative rating framing.

Furthermore, the RR × NRD interaction reached significance ($F(1,206) = 6.37$, $p < .01$) As predicted, the RR effect emerged as significant for low NRD condition but not for high NRD condition. LNRD (low negative rating density) consumers perceived more trust in the GRR (good rating ratio; M= 5.94) compared to the BRR (bad rating ratio; M=4.74). For HNRD (high negative rating density) consumers, there was no significant difference between the GRR and the BRR conditions (M = 3.25 vs. 2.55). Thus, H1 is supported by these findings. In addition, the rating framing did not interact with the RR manipulation; it interacted with the NRD manipulation ($F(1,206) = 4.41$, $p < .05$). The rating framing effect emerged as significant for low NRD condition but not for high NRD condition. LNRD (low negative rating density) consumers perceived a higher trust level in the positive rating faming condition

(M = 5.65) compared to the negative rating framing condition (M =5.07). For HNRD (high negative rating density) consumers, there was no significant difference between the two rating framing conditions (M = 2.98 vs. 2.84). H3 is verified.

In terms of the effect of risk perception, there is a main effect of RR($F(1, 206) = 97.40$, $p< .001$) and NRD($F(1, 206) = 503.67$, $p <.001$), such that the better the rating ratio, the lower the risk perception(M_{good} = 3.63 vs. M_{bad} = 4.79); the higher the negative rating density, the higher the risk perception(M_{high} = 5.52 vs. M_{low} = 2.97). Further, the H2 and H4 were tested. The results also revealed the expected interaction between RR and NRD ($F(1,206)=34.27$, $p < .001$). The RR effect emerged as significant for low NRD condition but not for high NRD condition. LNRD (low negative rating density) consumers perceived more risk perception than the BRR condition (M= 3.89) compared to the GRR condition (M = 2.09). For HNRD (high negative rating density) consumers, there was no significant difference between the BRR and the GRR condition (M = 5.76 vs. 5.30). There was also an interaction effect between rating framing and NRD on risk perception ($F(1,206) = 7.35$, $p < .01$). The rating framing effect emerged as significant for low NRD condition but not for high NRD condition. LNRD (low negative rating density) consumers perceived a lower perception level in the positive rating faming condition (M = 2.65) compared to the negative rating framing condition(M = 3.29). For HNRD (high negative rating density) consumers, there was no significant difference between the two rating framing conditions (M = 5.55 vs. 5.50). The rating framing did not interact with the RR manipulation.

DISCUSSION

The results of this study demonstrate that a greater rating ratio leads to higher trust of sellers, and that a higher density of negative rating leads to a lower level of trust. Positive rating framing is more significant than negative rating framing on trust. On the other hand, as predicated, a greater rating ratio creates a lower risk perception and a lower negative rating density creates a lower risk perception. For risky choice framing, the effect of more risk taking in the negative frame than in the positive frame is large and statistically significant. With regard to rating framing, risk perception under positive rating framing is lower than that under negative rating framing. A comparison of these three feedback resources shows that the negative rating density on trust and risk perception(η^2 = .77 and .71, respectively) is more influential than rating ratio (η^2 = .33 , .32, respectively) and rating framing g (η^2 = .024, .10, respectively).

Besides verifying main effects, it is also interesting to see the interaction effects among the three manipulated factors. The findings show that interaction effects occur in each situation except rating ratio and rating framing to trust and risk perception. The three-way interaction effect between rating ratio, negative rating density and rating framing has a minimal effect on trust and risk perception. However, two two-way interaction effects occurred on trust and risk perception. First, the negative rating density moderated the effect of rating ratio on trust and risk perception. For consumers in the low negative rating density condition, the effect of rating ratio will interfere with the level of trust and risk perception more significantly. It shows that the effects of the negative rating density and rating ratio may counteract each other. Second, positive rating framing gets higher trust and lower risk perception than negative rating framing, especially in the low negative rating density condition, but there is no difference in the two rating framing conditions on trust and risk perception when negative rating density is high. It is evident that rating framing moderated the effect of the negative density condition on trust and risk perception.

Trust is the extent to which one party is willing to depend on something or somebody in a given situation with a feeling of relative security, even though negative consequences are possible[18]. In the online auction environment, the buyer often has insufficient information about the seller. The buyer generally has no opportunity to see and try products, i.e. to “squeeze the oranges”, before he buys. The inefficiencies resulting from this information asymmetry can be mitigated through trust and feedback reputation. In feedback systems, information about the performance of a given seller is collected as ratings from other consumers who have had direct experience with that participant. This example shows that trust and risk taking plays a crucial role in computer mediated transactions and processes.

Feedback systems are typically based on public information in order to reflect the judgment of sellers. For example, eBay is a popular auction site that allows sellers to list items for sale, and buyers to bid for those items. The so-called Feedback Forum on e Bay gives buyer and seller the opportunity to rate each other as positive, neutral, or negative (i.e. 1, 0, -1). The Feedback Forum is a centralized feedback system where eBay collects all the ratings and computes the scores. The running total reputation score of each seller is the sum of positive ratings minus the sum of negative ratings. Furthermore, in order to provide more information about a seller’s recent record, the total of positive, neutral and negative ratings for the three different time windows (i) past 6 months, (ii) past month, and (iii) past 7 days are also displayed. For an overview of some empirical studies of eBay’s feedback system see Resnick, Zeckhauser, Swanson, & Lockwood (2006).

Our findings show that these types of feedback sources indeed influence consumers’ evaluation. Based on the framing effect, we aim to include new feedback to analyze how recommendations influence risk perception. The results show that simply presenting rating information with positive or negative framing will not create significant differences to risk perception. Positive or negative framing changes risk perception evaluation only when negative rating density is different. For example, with low negative rating density, negative rating framing will increase consumers’ risk perception. Among these three factors, negative rating density is the most influential to trust and risk perception, even higher than the rating ratio often used in the past. This paper also indicates it is more comprehensive to evaluate trust in online recommendations using multiple sources rather than only one. Prior research about rating systems, trust and risk perception focuses on a certain kind of rating ratio. It overlooks that in real circumstances other feedback sources affect trust and risk perception as well.

This paper finds out increasing trust towards recommendations does not necessarily translate into increasing risk perception; this would increase risk perception in some cases. As online auction markets mature, most consumers are becoming more accustomed to the system, and as scams increase, customers become more conservative towards recommendations. If customers notice a mismatch between sellers’ recommendations and self-expectation, this will increase their risk perception and discourage purchasing intention. We also suggest sellers pay more attention to how feedback influences consumers’ trust and risk perception in reaching their expected goals. In the near future, we also hope that the rich literature will consolidate around a set of feedback systems from sellers’ viewpoints and find the way into practical and online auction applications.

REFERENCES

[1] Ansari, A., Essegaier, S., Kohli, R. “Internet Recommendation Systems,”. Journal of Marketing Research, 2000, 37 (3), 363-375.

[2] Ba, S., Pavlou, P.A “Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior,”. MIS Quarterly, 2002, 26(3), 1-26.

[3] Ba, S. Whinston, A.B., Zhang, H. “Building Trust in Online Auction Markets through An Economic Incentive Mechanism,”. Decision Support Systems, 2003, 35(3), 273-286.

References available upon request.

THE POWER OF CHINA'S CHECKBOOK: IS SHANGHAI GOING TO BE AN INTERNATIONAL FINANCIAL CENTER?

Anindya K. Bhattacharya, Brooklyn College/CUNY, (212) 427-1733, anindyab@brooklyn.cuny.edu

ABSTRACT

This paper analyzes the policy options available to China for the utilization of its external reserves and examines alternatives uses for the reserves. It evaluates the national regulations/incentives that need to be implemented in order to establish an international financial center in Shanghai. It concludes that to the extent that hard currency funds are tied to onshore utilization by Chinese private smaller companies, an international financial center in Shanghai would contribute to China's economic development. However, the role to be played by such smaller enterprises in China's economic development process is a matter of a policy decision by the Chinese government.

Keywords: china; external reserves; shanghai; international financial center; diversification; feasibility **.**

INTRODUCTION

The huge foreign exchange reserves of China have attracted international attention. So far In 2009, the cumulative stock of China's external reserves has exceeded $2 trillion–the largest in the world, two-thirds of which were invested in U.S securities, overwhelmingly in Treasury securities. This phenomenon of excessive dependence on a single country for investment purposes is not in China's interests, and since China is the largest creditor to the United States providing half of the financing needed for the U.S. current account deficit, the financial dependence on China is also a matter of concern to the United States.

The topic is significant because Chinese options are limited with respect to external investments of foreign exchange reserves. China needs to hold hard currencies in its reserves because its own currency-renminbi (RMB)-is not convertible on the capital account and thus is not internationalized despite Chinese attempts to make it acceptable as a trading currency via some bilateral currency swap and trading arrangements. From the Chinese viewpoint, the surge in foreign exchange reserves has the undesirable twin effects of pushing up the value of RMB in foreign exchange markets and attracting "hot money" inflows. Moreover, Chinese investments in U.S. securities are generating minimal returns in view of low interest rates in the United States and are subject to both interest rate and exchange rate risks. China fears that massive deficit spending by the U.S. government to contain the economic-financial crisis might lead to inflation in the future and further dollar depreciation, thereby eroding the value of their dollar-denominated assets. China's recent attempts to hold more gold reserves, its proposal to expand the role of Special Drawing Rights as a "super-sovereign" reserve currency, and its plan to buy bonds to be issued by the International Monetary Fund (IMF) appear more to be expressions of concern about the "dollar trap" it is caught in (walking the fine line between diversification away from the dollar and maintaining the value of its dollar-denominated assets) than practical proposals that can be implemented by the international community.

The purpose of this paper is to (a) analyze the policy options available to China for the utilization of its foreign exchange reserves; and (b) examine alternative ways of putting these reserves to more fruitful uses that will serve Chinese interests. Specifically, the paper evaluates the feasibility of establishing an international financial center in Shanghai (as planned by China for 2020) so that Chinese small and medium-sized enterprises (SMEs) can have easier access to an institutionalized pool of hard currency financing. To this end, the paper explores the national regulations/incentives China needs to implement in order to make this new role for Shanghai a feasible proposition from the viewpoints of both the host country and the international financial community.

ANALYSIS OF CHINA'S POLICY OPTIONS

A Potpourri of Diversification Attempts: Chinese Overseas Investments

As long as China continues to have a substantial twin surplus in both the current account and the capital account, and build up huge savings (China's savings to GDP ratio is nearly 50 percent!), it will continue to accumulate vast foreign exchange reserves and be a net creditor to the world in general and the United States in particular. Since 2005, China has made several attempts to diversify away from the U.S. dollar. Such attempts included abandoning the fixed exchange rate regime between the RMB and the dollar, signing six bilateral currency swap agreements (with Malaysia, South Korea, Hong Kong, Belarus, Indonesia, and Argentina) using the Chinese currency as a trade settlement currency, and holding talks with Brazil concerning a bilateral agreement in which the Brazilian currency and the RMB would be used as trading currencies between the two countries. In 2009, the Governor of the People's Bank of China suggested a more extensive use of Special Drawing Rights as a 'super-sovereign" reserve currency to minimize having to rely on the currency of any single country. The Chinese have also expressed an interest in buying $20 billion worth of IMF bonds to be floated in public capital markets. In addition, they have been stockpiling commodities such as iron ore, crude oil, coal, aluminum, copper, nickel, tin, zinc, soybeans, grain, sugar and the like. China has also been unloading asset-backed securities of U.S. government-sponsored enterprises and buying more short-term U.S. Treasury bills than longer-term bonds .

In the 1990s, the Chinese government launched an initiative urging 100 targeted state-owned companies to go global. In 2006, it set up the China Investment Corporation-a sovereign wealth fund with funds of $200 billion-with a mandate to make cross-border investments. The problem with this strategy is that the Chinese companies generally lack international experience, brand name, scale and staying power– factors critical to success in the global arena. Few Chinese cross-border acquisitions have been profitable or have created value for their shareholders. In addition, Chinese investments abroad have run into political opposition, concerns about corporate governance and transparency, and uncertainties about the ultimate political objectives of the Chinese government.

These problems were illustrated by the bitter experience of Chinese investments abroad such as China National Overseas Oil Corporation's $18.5 billion attempt to purchase Unocal-a U.S. energy company; the bid by Haier, a Chinese electrical appliance company to acquire Maytag-a U.S. washing machine manufacturer; the purchase of French TV manufacturer Thomson; investments in Barclays of the United Kingdom and Fortis of Belgium and the Netherlands; the China Investment Corporation's injection of $3.3 billion in private equity group Blackstone Investment Fund and acquiring a 9 percent stake in Morgan Stanley. More recently, the $19.5 billion bid by Chinalco, a Chinese metals company, to acquire a majority stake in the Rio Tinto Mining Company of Australia ran into serious political opposition and was rejected.

The failure of the cross-border mega merger and acquisition deals involving substantial equity stakes means that henceforth Chinese companies will have to settle for deals overseas involving substantially smaller stakes in foreign companies, which are not "transformational" events and are not expected to make a dent in the attempt at diversifying its foreign portfolio. The flow of Chinese direct investment abroad accounted for only about 12.7 percent of China's total foreign exchange reserves in 2008, and the cumulative stock of Chinese direct investments abroad amounted to less than 6 percent of China's gross external financial assets at the end of 2008. Moreover, China's experience with respect to overseas investments in Africa, Central Asia and Latin America has been a mixed success story. Chinese investments in these resource-rich areas have certainly secured the procurement of critical raw materials, commodities, and energy sources such as oil and gas. But it has also created friction between the home and host governments, as illustrated by African concerns about the adverse impact on local industries and employment, and labor disputes in Chilean mines.

The Case For Establishing an International Financial Center in Shanghai

The option that seems to make the most sense for the Chinese is the establishment of an international financial center in Shanghai. China today has a credit market that functions well. Thanks to reforms instituted during the Asian financial crisis of the late 1990s, Chinese banks today are in much better shape than before (with the bad loan ratio declining from 50 percent to less than 10 percent.) They are relatively well capitalized and have fared better than western banks in coping with the global financial crisis. China also has the largest bank in the world (Industrial and Commercial Bank of China) in terms of market capitalization, assets, and deposits. Three out of China's four largest banks rank among the global top ten banks in market capitalization. Aside from the "power of the checkbook" stemming from its "savings glut" and huge foreign exchange reserves, China can also tap into the extensive experience of Hong Kong as an international financial center (particularly in the areas of liberal regulatory environment, low tax regime, and high-quality banking and infrastructure system) giving rise to the phenomenon known as "Shangkok" or Shanghai and Hong Kong combined together.

China has more than 40 million SMEs run by "emerging innovators and entrepreneurs" that generate most new jobs in the non-public sector and account for more than half of Chinese economic output. It has been argued that the real source of China's spectacular economic growth lies more in private entrepreneurs and "enterprises sponsored by townships and villages" (such as in Zhejiang province) than in state-controlled mega entities located in large urban centers. The SMEs will certainly need access to hard currency financing in order to succeed in a big way in the global market place. Historically, they have faced a great deal of difficulty in obtaining credit from either the credit market or the capital market due largely to high default risk, limited information, and lack of acceptable collateral. Despite government efforts to improve their access to bank and capital market financing, SME access to local currency financing, let alone foreign currency financing, remains severely constrained. As a result, the SMEs have no choice but to depend on internally-generated funds or the informal, gray market for their financing needs. Hence the case for additional sources of financing for SMEs. Having an institutionalized pool of foreign currencies at home in Shanghai will serve the financing needs of these smaller, entrepreneurial, countryside-based enterprises.

The Feasibility of establishing an International Financial Center in Shanghai

An international financial center can be defined as one that combines both onshore (providing financial services to residents and non-residents) and offshore (providing financial services to non-residents essentially in external currencies) activities. By definition, international financial centers must have large domestic economies, well-developed and liquid financial markets, and an adequate but minimal regulatory framework. Since Shanghai has the ambition of becoming a primary international financial center, the real issue in setting up an international financial center in China focuses on the net development impact on the host country after subtracting costs of setting up such an operation.

Benefits

An analysis of international financial centers in the Asia-Pacific region (such as Tokyo, Hong Kong and Singapore) shows that important benefits could be obtained by the host country such as local access to hard currencies, development of local financial skills and markets, and generation of local employment and government revenue. In Singapore, for example, international banks have made available to local and regional borrowers a pool of hard currency funds that have historically been used for development purposes. Aside from trade credit, Singaporean, Korean, and Japanese companies have made extensive use of the longer-term Asian dollar bond market as well as the secondary market for negotiable certificates of deposit.

As mentioned before, having an international financial center in Shanghai could fulfill a substantial part of the funding needs of private-sector SMEs that might otherwise find it difficult to obtain hard currency financing. To the extent that hard currency funds are tied to onshore utilization by SMEs, establishing an international financial center in Shanghai would contribute to China's economic development by further improving the export competitiveness of its products. However, the role to be played by SMEs in China's economic development process is a matter of a policy decision by the Chinese government.

Costs

For this internationalization strategy to work, China needs to undertake several drastic steps that will involve an evaluation of the costs involved. An analysis of the experience of other international financial centers shows that the host country must offer a package of "minimal regulations/maximum incentives", chief among which are the following :

- the abolition of liquidity requirements for foreign currency deposits;
- the abolition of interest-withholding taxes on foreign currency transactions;
- lowering corporate income taxes on profits of international banks and financial institutions;
- liberalizing exchange control to permit international bank profits to be remitted freely;
- making the local currency convertible on the capital account to ensure free movement of capital in and out of the country;
- developing the capital market (specifically the bond market) and keeping a portion of that market for use by foreigners; and
- maintaining a laissez-faire regulatory environment, improving the legal, professional and support services infrastructure, and ensuring the supply of skilled labor.

Definitely, there are risks associated with this strategy. As the experience of Hong Kong during the Asian financial crisis of the late 1990s showed, the island was deluged with a massive influx of foreign currencies, which were then swapped for local currency for speculative investments in local stock and real estate markets. This "double play" in the money market and the stock market by institutional investors resulted in extraordinary intervention in the stock market by the Hong Kong Monetary Authority, which had historically been known for its laissez-faire approach to the market place. Due mainly to expectations of continued RMB appreciation against the dollar, China today is inundated with inflows of substantial speculative capital or "hot money", the cumulative stock of which is estimated to be between $500 billion and $1.75 trillion. The Asian financial crisis showed that foreign speculative portfolio capital can flow out of a country just as easily as it can come into the same country. A related significant cost of establishing an international financial center at home lies in capital outflows initiated by residents, especially when local conditions are difficult for them.

EVALUATION OF STEPS UNDERTAKEN BY CHINA TO DEVELOP SHANGHAI AS AN INTERNATIONAL FINANCIAL CENTER

China has been liberalizing its financial sector ever since 1992-1993. Plans for establishing an international financial center in Shanghai are, however, of recent origin. China senses an opportunity for Shanghai in the global financial crisis which has tarnished the reputation of London and New York. Still, in the typical cautionary and evolutionary mode, it wants to accomplish this goal over decades rather than in a drastic way all at once. It has taken several half-hearted steps but still has not gone all the way to undertake the kinds of sweeping measures that are needed if Shanghai is to succeed in its plan to join the triad of primary financial centers.

Thus, for example, China plans to allow qualified foreign companies to list on the Shanghai stock exchange and to invest in RMB-denominated shares and bonds under a quota but has no immediate plans for making its currency fully convertible (for fear of capital outflows)–a phenomenon that will certainly stand in the way of profit repatriation by foreign companies.

Secondly, China has issued a draft list of rules for the proposed Growth Enterprise Market (GEM) in Shenzen–a NASDAQ-style over-the-counter market-that will expand the financing options for SMEs, particularly start-up technology firms, but the credit rating agencies have yet to develop a comprehensive credit monitoring and rating system for all SMEs.

Third, China is actively developing futures trading in essential commodities such as copper, corn, soybean and wheat as well as foreign exchange forwards and swaps but has not taken any steps yet in expanding the role of derivatives trading beyond commodities and foreign exchange.

Fourth, since June 2006 China has been allowing residents to invest overseas but has not actively eliminated the restrictions on investments in foreign equity by domestic investors. The accumulated stock of Chinese investments in foreign securities was a meager $20.8 or less than 1 percent of China's gross total external assets in 2008.

Fifth, China has taken an important step in reforming its equity market by allowing formerly non tradable shares to be fully tradable by 2012. Hopefully, this measure will supply liquidity into the domestic equity market, enlarge the base of domestic institutional as well as retail investors, and reduce the reliance on bank financing. However, key accompanying measures to bolster the equity market – strengthening the accounting and auditing standards, improving the legal framework in the areas of enforcement and support, upgrading the transparency of corporate governance, and changing the managerial mind set from growth to profitability–are still lacking or are inconsistently applied.

Sixth, China has taken several steps to promote the development of its bond market, especially the inter-bank bond market. Such steps include increasing the number of market participants, allowing commercial banks to issue subordinated bonds, encouraging international development institutions to issue RMB bonds in the domestic market, allowing foreign banks to trade and underwrite corporate bonds, and introducing bonds with floating interest rates and interest rate derivatives transactions. But the Chinese bond market is still heavily dominated by the government and government-owned policy banks, and the proportion of direct financing raised in the bond market by the corporate sector is miniscule due to low liquidity, absence of diversified financial products and trading instruments, and lack of depth and breadth of the market. Commercial banks continue to remain the most active investors holding over one-third of corporate bonds, and the involvement of the non-financial corporate sector is minimal, amounting to only 6 percent. Moreover, the secondary market for corporate bonds is not yet fully developed since most bonds are traded in the interbank market.

Lastly, China is in the rudimentary stage of developing the financial infrastructure, including recruitment and retention of domestic and expatriate financial talent. Currently there are only 40 qualified foreign institutional investors with approved quotas in China. Shanghai has attempted to cope with the shortage of workers by letting in "migrant workers" from other parts of China but strict immigration controls are still retained on imported foreign labor. China needs to build a qualified workforce, attract financial talent from all over the world, and develop Shanghai as a hub for cutting-edge, high value-added financial areas such as financial engineering, risk management, information technology, and actuarial science.

CONCLUSION

As long as China maintains a twin surplus in its balance of payments and enjoys a high savings rate, it will continue to build up enormous foreign exchange reserves. In the absence of fundamental macroeconomic adjustments (such as letting its currency float freely and promoting higher domestic consumption), Chinese choices are limited with respect to external investment possibilities of these reserves. Its current foreign portfolio is heavily tilted towards dollar-denominated assets–a situation that is not in China's interests. A better way of using these reserves would be to develop Shanghai as an international financial center and institutionalize the dollars there to be tapped by Chinese SMEs.

However, for Shanghai to play the role of an international financial center, China would have to make its currency convertible on the capital account, relax its control over the foreign exchange market, allow foreign institutional investors to invest in more types of RMB-denominated assets and financial derivative products, develop its bond market for active use by non-state as well as foreign participants, and develop the financial services infrastructure, including education and training for finance and related professionals. Such drastic changes do not mesh well with the current Chinese cautious and evolutionary approach to financial market liberalization. While radical financial reform cannot be expected to take place overnight in an emerging market like China, converting Shanghai into a full-fledged international financial center is decades away under the current way of doing business in China.

REFERENCES

Available upon request from Anindya K. Bhattacharya.

Corporate Financing Decisions and Investment Inefficiency

Hsiao-Fen Hsiao, Mingdao University, 369 Wen-Hua Rd., Peetow, Changhua 52345, Taiwan,Tel: +886-4-887-6660#7527, fen@mdu.edu.tw
Chun-An Li, National Yunlin University of Science and Technology, 123 University Rd., Sec. 3, Douliou , Yunlin 64002, Taiwan, Tel: +886-5-534-2601#5411, liica@yuntech.edu.tw

ABSTRACT

We set out in this study to examine the relationship between the financing decisions of firms and their level of inefficient under- or over-investment. We carry out two types of empirical tests, analyzing the level of investment for maximum firm value (dependent upon the quality of the investment opportunities) so as to gauge the relative magnitude of the under-investment and over-investment effects. Firstly, from an examination of the relative changes in the investment intensity of firms following public equity and debt offerings, we find that those firms issuing equity tend to increase (reduce) their investment by less (more) than similar firms issuing debt. Secondly, we examine different financing decisions in an attempt to determine whether such decisions represent the underlying cause of such inefficient investment.

Keywords: Financing decisions; Investment Inefficiencies; investment opportunities.

INTRODUCTION

There are numerous studies within the recent corporate finance literature discussing the relationship between investment strategies and firm value (Heaton, 2002; Bertrand and Mullainathan, 2003; Pinkowitz and Williamson, 2005 and Faulkender and Wang, 2006). Their results reveal that managers may often be found to be prone to under-investment (over-investment) by engaging in less than (more than) the optimum level of investment. The focus in this study is therefore on the influences that corporate financing decisions have on the under-investment or over-investment behavior of firms.

Conversely, many other studies within the financial literature have demonstrated numerous examples of over-investment (where a firm's actual investment is higher than the optimum level). Jensen (1986), for example, suggests that negative NPV projects provide managers with incentives to use their free cash flow, a situation which would not occur if they were required to raise their capital externally at higher costs. In other words, fluctuations in free cash flow can give rise to over-investment behavior. Managers will find incentives to over-invest because of the non-pecuniary benefits associated with larger firms (Jensen, 1986; Stulz, 1990).

In other words, the 'dual issue' versus 'equity issue' regressions enable us to examine the effects of market performance on the choices relating to the form of financing whilst holding market timing constant. Therefore, any differences in market performance observed between dual issuers and equity issuers can be attributed to the trade-off hypothesis.

Our results show that in many cases where firms have valuable investment opportunities, different financing decisions can increase the over-investment problem, whereas in many other cases where firms have no real valuable investment opportunities, different financing decisions can increase the under-investment problem. Thus, in order to increase the value of their firm, such managers need to revise their level of investment in an attempt to achieve the optimum level.

THE MODEL

Management Investment Decisions

Although the extant literature provides evidence of the impact of corporate financing decisions on the level of corporate investment, it does not reveal the level of efficiency associated with such investment, a factor which may be due partly to the difficulties involved in measuring the optimum level of investment. The only determinants of optimal investment decisions, therefore, we classify firms into two groups, those where Tobin's q is greater than 1(hereafter, VP firms), and all other firms, which are regarded as non-valuable project firms (hereafter, NVP firms). In the present study, we follow Morgado and Pindado (2003) to determine whether the relationship between firm value and investment is quadratic, which would thereby imply an optimum level of investment, depending upon the quality of their investment opportunities. Model I is described as follows:

$$\begin{aligned}\frac{V_{i,t}}{K_{i,t-1}} = \beta_0 &+ (\beta_1 + \gamma_{11} G_{i,t}(Q_{i,t}-1) + \gamma_{12} H_{i,t}(Q_{i,t}-1))\left(\frac{I_{i,t}}{K_{i,t-1}}\right) \\ &+ (\beta_2 + \gamma_{21} G_{i,t}(Q_{i,t}-1) + \gamma_{22} H_{i,t}(Q_{i,t}-1))\left(\frac{I_{i,t}}{K_{i,t-1}}\right)^2 \\ &+ \beta_3\left(\frac{\Delta B_{i,t}}{K_{i,t-1}}\right) + \beta_4\left(\frac{D_{i,t}}{K_{i,t-1}}\right) + \beta_5 ROA_{i,t} + e_{i,t}\end{aligned} \tag{1}$$

where $V_{i,t}$ is the market value of the shares of firm i at the end of period t; $I_{i,t}$ is the investment undertaken by firm i in period t; $\Delta B_{i,t}$ is the increment in the market value of long-term debt; $D_{i,t}$ is the dividend paid in period t; $K_{i,t-1}$ is the replacement value of the assets at the end of period $t-1$; and $ROA_{i,t}$ is the return on assets in period t. We also define a dummy variable for each firm: $Q_{i,t}$ is the value of *Tobin's q* for firm i at the end of period t; $G_{i,t}$ is equal to 1 for those firms with a *Tobin's q* value greater than 1 during the period, otherwise 0; $H_{i,t}$ is equal to 1 for those firms with a *Tobin's q* value which is less than 1 during the period, otherwise 0.

Consequently, following the estimation of the model, if we differentiate the firm value variable with regard to the investment variable, with the first derivative being equal to 0, and solving for the investment variable, we can obtain:

$$\left(\frac{I_{i,t}}{K_{i,t-1}}\right)^* = -\frac{\beta_1 + \gamma_{11} G_{i,t}(Q_{i,t}-1) + \gamma_{12} H_{i,t}(Q_{i,t}-1)}{2(\beta_2 + \gamma_{21} G_{i,t}(Q_{i,t}-1) + \gamma_{22} H_{i,t}(Q_{i,t}-1))}$$

$$\Rightarrow \begin{cases} \left(\dfrac{I_{i,t}}{K_{i,t-1}}\right)^*_{VP_{i,t}} = -\dfrac{\beta_1 + \gamma_{11}(Q_{i,t}-1)}{2(\beta_2 + \gamma_{21}(Q_{i,t}-1))} & \textit{if } Q_{i,t} > 1, \textit{ then } \text{ VP firms} \\ \left(\dfrac{I_{i,t}}{K_{i,t-1}}\right)^*_{NVP_{i,t}} = -\dfrac{\beta_1 + \gamma_{12}(Q_{i,t}-1)}{2(\beta_2 + \gamma_{22}(Q_{i,t}-1))} & \textit{if } Q_{i,t} < 1, \textit{ then } \text{ NVP firms} \end{cases} \tag{2}$$

Finally, if the second partial derivative of the firm value variable (relating to the investment variable) is negative. As a result, we propose the following hypothesis:

Hypothesis 1: *For maximized firm value, the optimal investment level of firms with valuable investment opportunities (VP), which will be larger than the optimal investment level of firms with non-valuable investment opportunities (NVP).*

Investment Decisions and Corporate Financing Decisions

The financing decisions of a firm will clearly affect the optimal level of investment; this will in turn differ for each firm, depending upon the quality of the investment opportunities. That leads to Model II, which is described as follows:

$$\begin{aligned}\frac{V_{i,t}}{K_{i,t-1}} = \beta_0 &+ [\beta_1 + \delta_{11} G_{i,t}(Q_{i,t}-1) + \delta_{12} H_{i,t}(Q_{i,t}-1) + \alpha_{11} E_{i,t} + \alpha_{12} D_{i,t} + \alpha_{13} Du_{i,t}]\left(\frac{I_{i,t}}{K_{i,t-1}}\right) \\ &+ [\beta_2 + \delta_{21} G_{i,t}(Q_{i,t}-1) + \delta_{22} H_{i,t}(Q_{i,t}-1) + \alpha_{21} E_{i,t} + \alpha_{22} D_{i,t} + \alpha_{23} Du_{i,t}]\left(\frac{I_{i,t}}{K_{i,t-1}}\right)^2 \\ &+ \beta_3\left(\frac{\Delta B_{i,t}}{K_{i,t-1}}\right) + \beta_4\left(\frac{D_{i,t}}{K_{i,t-1}}\right) + \beta_5 ROA_{i,t} + e_{i,t}\end{aligned} \tag{3}$$

where $E_{i,t}$ is a dummy variable which is equal to 1 if the corporate financing decision involves equity issuers, otherwise 0; $D_{i,t}$ is a dummy variable which is equal to 1 if the corporate financing decision involves debt issuers, otherwise 0; $Du_{i,t}$ is a dummy variable which is equal to 1 if the corporate financing decision involves dual issuers, otherwise 0.

The model defined in Equation (3) relates to firm value and investment, with additional controls for corporate financing decisions which are the other main factors of investment. From the maximum in Equation (3), corporate financing decisions are the level of investment for VP firms:

$$\left(\frac{I_{i,t}}{K_{i,t-1}}\right)^{**} = -\frac{[\beta_1 + \delta_{11} G_{i,t}(Q_{i,t}-1) + \delta_{12} H_{i,t}(Q_{i,t}-1) + \alpha_{11} E_{i,t} + \alpha_{12} D_{i,t} + \alpha_{13} Du_{i,t}]}{2[\beta_2 + \delta_{21} G_{i,t}(Q_{i,t}-1) + \delta_{22} H_{i,t}(Q_{i,t}-1) + \alpha_{21} E_{i,t} + \alpha_{22} D_{i,t} + \alpha_{23} Du_{i,t}]} \quad (4)$$

$$\Rightarrow \begin{cases} \left(\frac{I_{i,t}}{K_{i,t-1}}\right)^{**}_{VP_{i,t}} = -\frac{\beta_1 + \delta_{11}(Q_{i,t}-1) + \alpha_{11}}{2(\beta_2 + \delta_{21}(Q_{i,t}-1) + \alpha_{21})} & if \ E_{i,t} = 1, \ then \ \text{equity issuers} \\ \left(\frac{I_{i,t}}{K_{i,t-1}}\right)^{**}_{VP_{i,t}} = -\frac{\beta_1 + \delta_{11}(Q_{i,t}-1) + \alpha_{12}}{2(\beta_2 + \delta_{21}(Q_{i,t}-1) + \alpha_{22})} & if \ D_{i,t} = 1, \ then \ \text{debt issuers} \\ \left(\frac{I_{i,t}}{K_{i,t-1}}\right)^{**}_{VP_{i,t}} = -\frac{\beta_1 + \delta_{11}(Q_{i,t}-1) + \alpha_{13}}{2(\beta_2 + \delta_{21}(Q_{i,t}-1) + \alpha_{23})} & if \ Du_{i,t} = 1, \ then \ \text{dual issuers} \end{cases}$$

and the level of investment for NVP firms:

$$\left(\frac{I_{i,t}}{K_{i,t-1}}\right)^{**} = -\frac{[\beta_1 + \delta_{11} G_{i,t}(Q_{i,t}-1) + \delta_{12} H_{i,t}(Q_{i,t}-1) + \alpha_{11} E_{i,t} + \alpha_{12} D_{i,t} + \alpha_{13} Du_{i,t}]}{2[\beta_2 + \delta_{21} G_{i,t}(Q_{i,t}-1) + \delta_{22} H_{i,t}(Q_{i,t}-1) + \alpha_{21} E_{i,t} + \alpha_{22} D_{i,t} + \alpha_{23} Du_{i,t}]} \quad (5)$$

$$\Rightarrow \begin{cases} \left(\frac{I_{i,t}}{K_{i,t-1}}\right)^{**}_{NVP_{i,t}} = -\frac{\beta_1 + \delta_{12}(Q_{i,t}-1) + \alpha_{11}}{2(\beta_2 + \delta_{22}(Q_{i,t}-1) + \alpha_{21})} & if \ E_{i,t} = 1, \ then \ \text{equity issuers} \\ \left(\frac{I_{i,t}}{K_{i,t-1}}\right)^{**}_{NVP_{i,t}} = -\frac{\beta_1 + \delta_{12}(Q_{i,t}-1) + \alpha_{12}}{2(\beta_2 + \delta_{22}(Q_{i,t}-1) + \alpha_{22})} & if \ D_{i,t} = 1, \ then \ \text{debt issuers} \\ \left(\frac{I_{i,t}}{K_{i,t-1}}\right)^{**}_{NVP_{i,t}} = -\frac{\beta_1 + \delta_{12}(Q_{i,t}-1) + \alpha_{13}}{2(\beta_2 + \delta_{22}(Q_{i,t}-1) + \alpha_{23})} & if \ Du_{i,t} = 1, \ then \ \text{dual issuers} \end{cases}$$

Investment Inefficiency Measures

Since information asymmetry between managers and shareholders leads to inefficient investment, we empirically investigate the relative magnitudes of the under- and over-investment of firms caused by corporate financing decisions.

Investment inefficiency caused by equity (debt or dual) financing for VP firms:

$$\left(\frac{I_{i,t}}{K_{i,t-1}}\right)^{**}_{VP_{i,t}} - \left(\frac{I_{i,t}}{K_{i,t-1}}\right)^{*}_{VP_{i,t}}$$
$$= -\frac{\beta_1 + \delta_{11}(Q_{i,t}-1) + \alpha_{1k}}{2(\beta_2 + \delta_{21}(Q_{i,t}-1) + \alpha_{2k})} - \left(-\frac{\beta_1 + \gamma_{11}(Q_{i,t}-1)}{2(\beta_2 + \gamma_{21}(Q_{i,t}-1))}\right) = \begin{cases} > 0 & \text{overinvestment} \\ < 0 & \text{underinvestment} \end{cases}$$
$$k = 1: equity\ ; 2: debt\ ; 3: dual \quad (6)$$

Investment inefficiency caused by equity (debt or dual) financing for NVP firms:

$$\left(\frac{I_{i,t}}{K_{i,t-1}}\right)^{**}_{NVP_{i,t}} - \left(\frac{I_{i,t}}{K_{i,t-1}}\right)^{*}_{NVP_{i,t}}$$
$$= -\frac{\beta_1 + \delta_{12}(Q_{i,t}-1) + \alpha_{1k}}{2(\beta_2 + \delta_{22}(Q_{i,t}-1) + \alpha_{2k})} - \left(-\frac{\beta_1 + \gamma_{12}(Q_{i,t}-1)}{2(\beta_2 + \gamma_{22}(Q_{i,t}-1))}\right) = \begin{cases} > 0 & \text{overinvestment} \\ < 0 & \text{underinvestment} \end{cases}$$
$$k = 1: equity\ ; 2: debt\ ; \ 3: dual \quad (7)$$

As a result, we propose the following hypothesis:

Hypothesis 2: *The firms with valuable investment opportunities (VP), will tend to overestimate project value, debt issuers are therefore more likely to demonstrate a phenomenon of over-investment than dual issuers.*

Hypothesis 3: T*he firms with non-valuable investment opportunities (NVP), will tend to underestimate project value, there will tend to be the greater likelihood of equity issuers exhibiting a phenomenon of over-investment than both debt issuers and dual issuers.*

EMPIRICAL EVIDENCE

Data Sources

Following MacKie-Mason (1990) and Hovakimian et al. (2001), we identify security issuing firms using annual firm-level data from the non-financial quoted companies in Taiwan, with the primary data source being the Taiwan Economic Journal (TEJ) database. Our panel dataset was constructed covering the period from 1998 to 2008 in order to avoid endogeneity and unobservable heterogeneity; that is, an unbalanced panel comprising of 1,449 companies during the sample period.

Results

The results of our examination of the relationship existing between firm value and investment which dependent upon the quality of investment opportunities. The results of the estimates of Model I are provided in Table 1.

Table 1

Estimation of the manager's investment decisions model using asymmetric panel data methodology to avoid endogeneity and heterogeneity.

	β_0	β_1	β_2	β_3	β_4	β_5
Coefficient	0.5236 (0.0145)***	-0.8022 (0.0143)***	0.0143 (0.0062)**	0.5842 (0.0745)	9.1437 (0.0545)***	0.0192 (0.0008)***
	γ_{11}	γ_{12}	γ_{21}	γ_{22}		
Coefficient	12.331 (0.1734)***	-0.6085 (0.1297)***	-12.4412 (0.3665)***	0.0130 (0.0063)**		

Note: standard errors in ( ). *:10%, **:5%, ***:1% significance level

By maximizing firm value, the optimum level of investment for VP and NVP firms will differ, dependent upon the quality of investment opportunities. Therefore, support for Hypothesis 1. The coefficient for the 'increment in debt' variable is 0.5842, with insignificance at the 10% level, whilst the 'dividends' variable is 0.1437, also with significance at the 1% level. Hence, as stipulated in the capital-market arbitrage condition, current dividends are a source of value creation for shareholders.

Table 2

Estimation of corporate financing decisions and investment decisions model using asymmetric panel data

	β_0	β_1	β_2	β_3	β_4	β_5
Coefficient	0.5177 (0.0145)***	-0.5557 (0.1516)***	-0.0280 (0.0106)***	0.7203 (0.0812)	0.2512 (0.0583)***	0.0188 (0.0008)***
	δ_{11}	δ_{12}	α_{11}	α_{12}	α_{13}	δ_{21}
Coefficient	12.3466 (0.1781)***	-0.2232 (0.1671)	2.5815 (0.7914)***	-0.1191 (0.0929)	2.5690 (1.1642)**	-12.8371 (0.3771)***
	δ_{22}	α_{21}	α_{22}	α_{23}		
Coefficient	-0.0312 (0.0109)***	2.5690 (1.1642)**	0.0196 (0046)***	-6.2288 (2.6796)**		

Note: standard errors in ( ). *:10%, **:5%, ***:1% significance level

The results of the estimates of Model II are provided in Table 2, from which we can see that the coefficients for the equity issuers are –2.5815 for α_{11} and 9.0222 for α_{21}, both with significance at the 1% level. The coefficients for the debt issuers are –0.1191 for α_{12}, with insignificance, and 0.0196 for α_{22}, with significance at the 1% level.

The coefficients for the dual issuers are 2.5690 for α_{13} and –6.2288 for α_{23}, both with significance at the 5% level. Since β_1 is found to be –0.5557 and β_2 is –0.0280, when combining these results with the coefficients of 12.3466 for δ_{11} and –12.8371 for δ_{21}, we can confirm the existence of a quadratic relationship between firm value and investment for VP firms. Furthermore, given that the coefficient on δ_{12} is –0.2232, with no significance, whilst that on δ_{22} is –0.0312, with significance at the 1% level, these findings also enable us to confirm the same quadratic relationship for NVP firms.

Table 3

Estimation of investment dependent upon the quality of the investment opportunities and corporate financing decisions affects level of under-/overinvestment.

	VP firms	NVP firms
Panel A Equity issuers(488 observations)		
Over-investment	18.03%	0
Under-vestment	3.28%	85.87%
Panel B Debt issuers(6067 observations)		
Over-investment	26.77%	7.12%
Under-vestment	1.60%	37.41%
Panel C Dual issuers(192 observations)		
Over-investment	7.81%	0
Under-vestment	32.81%	84.90%

Whether the corporate financing decisions taken by a firm will affect their investment strategies will be heavily dependent upon the quality of their investment opportunities. The results of the estimates of inefficient levels of (under/over) investment are provided in Table 3. We find that the existence of a phenomenon of over-investment is largely dependent upon the valuable investment opportunities that are available to equity or debt issuers, but that this is not the case for dual issuers. Conversely, whether any of the firms will exhibit a phenomenon of under-investment will be largely dependent upon the presence of non-valuable investment opportunities.

Table 4

Test corporate financing decisions affects level of under-/overinvestment difference between VP firms and NVP firms.

	Mean	Standard deviation	Wilcoxon rank sum test
Panel A: VP firms			
Equity issuers(488 observations)	-1.2773	0.9770	

Debt issuers(6067 observations)	-0.1705	0.1044	
Dual issuers(192 observations)	0.1900	0.4228	
Z (H_0:same population distribute in Equity issuers and debt issuers)			-14.055***
P(overinvestment phenomenon in Equity issuers more than in debt issuers)			0.309
Z (H_0:same population distribute in Equity issuers and dual issuers)			-1.571
P(overinvestment phenomenon in Equity issuers more than in dual issuers)			0.461
Z (H_0:same population distribute in debt issuers and dual issuers)			12.194***
P(overinvestment phenomenon in debt issuers more than in dual issuers)			0.758***
Panel B: NVP firms			
Equity issuers(488 observations)	-27.9955	0.1559	
Debt issuers(6067 observations)	10.8527	33.7657	
Dual issuers(192 observations)	-27.7443	0.1666	
Z (H_0:same population distribute in Equity issuers and debt issuers)			7.453***
P(overinvestment phenomenon in Equity issuers more than in debt issuers)			0.601**
Z (H_0:same population distribute in Equity issuers and dual issuers)			0.423
P(overinvestment phenomenon in Equity issuers more than in dual issuers)			0.51*
Z (H_0:same population distribute in debt issuers and dual issuers)			-4.978***
P(overinvestment phenomenon in debt issuers more than in dual issuers)			0.395

Note:*:10%, **:5%, ***:1% significance level

As we can see from Table 4, corporate financing decisions do have an effect on inefficiency levels in firm investment. Since VP firms will tend to overestimate project value, debt issuers are therefore significantly more likely to demonstrate a phenomenon of over-investment than dual issuers, with significance. Therefore, support for Hypothesis 2. Conversely, NVP firms will tend to underestimate project value, and as such, there will tend to be the greater likelihood of equity issuers exhibiting a phenomenon of over-investment than both debt issuers and dual issuers, again with significance. Therefore, support for Hypothesis 3.

CONCLUSIONS

We set out in this study to examine the relationships that exist between corporate financing decisions and the extent of inefficient under-/over-investment amongst a sample of Taiwanese firms. The findings of this study make two fundamental contributions to the further understanding of investment policy decisions. We present evidence to show that there is a tendency amongst Taiwanese firms for both equity and debt issuers to over-invest, and that this is largely dependent upon the valuable investment opportunities available to them. We also find a tendency for under-investment which is more acute for corporate financing decisions based upon non-valuable investment opportunities, than for such decisions based upon valuable investment opportunities.

References available upon request from Corresponding Author: Hsiao-Fen Hsiao fen@mdu.edu.tw

Mispricing and Price Momentum

Chun-An, Li , 05-5342601-2601, Email：liica@yuntech.edu.tw
Chih-Cheng, Yeh, 04-23892088-3512, Email：g9420802@yuntech.edu.tw
Ling-Hsuan, Huang, 07-6011000-1203, Email：virginia3030@yahoo.com.tw
Kun-Mei Pan 04-22196022, Email：kmpan@ntit.edu.tw

ABSTRACT

The paper examines whether there exists difference between the magnitude of mispricing of high priced stocks and that of low priced ones, and if such difference has any impact on investment performance. Based on the difference, momentum strategies are constructed to observe if they generate any extraordinary returns. Empirical results show that low priced stocks yield higher returns because their misrpricing phenomenon is significantly lower that that of high priced stocks. Returns of four momentum strategies increases with an increase in the holding period, indicating that constructing a momentum portfolio based on the magnitude of mispricing is an effective strategy.

Keyword: mispricing, momentum, low priced

INTRODUCTION

Since Mackinlay and Ramaswamy(1988) established the mispricing model, researchers have proposed several methods to measure the degree of deviation of the current stock price from its fundamental value and found that value stocks earn higher returns than growth stocks(Laknoishok et al.,1994; Frankel and Lee ,1998). Most subsequent researchers have explained the mispricing phenomenon from the perspective of risk and investor behavioral pitfalls(Skinner and Sloan, 2002; Chen et al., 2008). There are also researchers believing that insider trading is correlated to mispricing and even the magnitude of mispricing can be used as a signal to determine whether the company is making any capital decision-making(Kahle, 2000；Graham and Harvey, 2001). Insiders who have better knowledge about the company will be induced to buy value stocks that are underpriced and sell growth stocks that are overpriced to gain profit(Rozeff and Zaman(1998)；Laknoishok and Lee, 2001). Thus, the mispricing phenomenon of stock price obviously offers a sign that the company is undertaking capital decision-makings and a signal for possible hidden momentum profits.

Generally speaking, the high share price reflects the company is being operated soundly and investors have an optimistic view on the future prospects of the firm. It is often not a reflection of its true value. On the contrary, a low share price reflects the company is not operated soundly and investors have no confidence in the future prospects of the firm, leading to a share price that is lower than its true value(Blitzer and Dash, 2002).Whether such situation will result in a more severe mispricing for high priced stocks than low priced stocks? Jenter(2005) follows Book-to-Market Ratio(B/M Ratio) to identify if the share price of a firm is being overpriced or underpriced. A high B/M ratio is a sign that that type of stock is being underpriced relative to its fundamental value while a low B/M ratio refers that stocks are being overpriced against the fundamental value. In the way just indicated, there will be four clusters of stocks in the market, namely, high priced stocks that are overpriced(high-priced/overpriced cluster), high priced stocks that are underpriced(high-priced/underpriced cluster), low priced stocks that are overpriced(low-priced/overpriced cluster), and low priced stocks that are underpriced(low-priced/overpriced cluster). A high mispricing often refers to higher degree of deviation from its fundamental value. Then, there must be an extraordinary momentum profit when buying underpriced stocks with low magnitude of mispricing and selling overpriced stocks with high magnitude of mispricing.

Taiwan stock market investors comprise a large proportion(70% or more) of individual ones, and compared to the developed countries, the mispricing phenomenon is more apparent, making it a suitable study subject for the paper. The purposes of the study include the following:

1. To identify whether there is a difference between the magnitude of mispricing of high priced stocks and that of low priced stocks;

2. To establish momentum strategies based on the difference between the magnitude of mispricing of high priced stocks and that of low priced stocks to observe if there are any significantly extraordinary returns;

3. To investigate if factors such as firm size, January effect, market status, and industry have an influence on momentum profits.

TESTS AND RESULTS

Samples of the study were selected from the companies listed on the Taiwan Stock Exchange and their performance in the period from January 2001 through June 2007 was investigated based on data collected from TEJ. A stock whose price exceeds its par value every month and is in the top 20% is called a high priced stock, and that whose price is below its par value every month is a low priced stock. Meanwhile, the B/M Ratio is employed to identify if the share price of a firm is being overpriced or underpriced. When the B/M Ratio is low, the stock price tends to be overpriced; on the contrary, the stock price has a greater tendency to be underpriced when the B/M Ratio is high. Sorting stocks in order of increasing BM Ratio, those in the top 20% were categorized as the overpriced cluster and those in the bottom 20% were classified into the underpriced cluster. Total four clusters were obtained: high-priced/overpriced cluster, high-priced/underpriced cluster, low-priced/overpriced cluster, low-priced/overpriced cluster. The return was calculated with the weights given by the market value. In this paper, the formation period was one month and the holding period was 1 to 12 months. Under the buy and hold strategy, the paper calculated the cumulative returns of these four investment portfolios during the holding period and observed the magnitude of mispricing.

Through the characteristics of B/M Ratio, the study observed the magnitude of mispricing of high priced stocks(overpriced & underpriced) and low priced stocks(overpriced & underpriced) . Table 1 shows the average B/M Ratio for high priced stocks is 0.129, while that for low priced stocks is 0.805, indicating that growth stocks are generally high priced and value stocks are usually low priced. Meanwhile, high-priced/overpriced clusters deviate from the average market price by 334%, which is higher than 242% of low-priced/overpriced clusters; high-priced/underpriced clusters deviate from the average market price by 481%, which is higher than 303% of low-priced/underpriced clusters. This indicates high priced stocks have a more severe mispricing than low priced stocks. Furthermore, when there is a holding period of 12 months, low-priced/overpriced clusters and low- priced/underpriced clusters perform significantly better than high-priced/overpriced clusters and high-priced/underpriced clusters, indicating low priced stocks have higher returns. The empirical results support the arguments by Loughran and Ritter (1996) that low price effects do exist in the stock market, and they are obviously related to the low magnitude of mispricing of low priced stocks.

Table 1. Characteristic of B/M ratio for High-price and Low-price

	Mean			All/Over	Under/All	Cumulative Return	
	Over	All	Under			Over	Under
High-price	0.039	0.129	0.62	331%	481%	0.110	0.061
Low-price	0.333	0.805	2.439	242%	303%	0.332	0.440

Now that high priced stocks have a larger magnitude of mispricing than low priced stocks, then, what effects does this have on the return of investment portfolio? The empirical results indicate that, except the high-priced/underpriced cluster, all the other three portfolios are performing better than market returns. Besides, due to the low magnitude of mispricing of low priced stocks, the difference between the return of the low-priced/overpriced cluster and that of the low-priced/overpriced cluster is not getting bigger with the increase of holding period(2.9%→ 2.7%); however, the difference between the return of the high-priced/overpriced cluster and that of the high-priced/underpriced cluster does expand with the increase of holding period(6.1%→ 20.6%), causing the difference between the return of the low priced stocks(overpriced & underpriced) and that of the high-priced stocks(overpriced & underpriced) to become bigger. With the 12-month holding period, the difference between the return of low-priced/overpriced clusters and that of high-priced/overpriced clusters is (1.18%→ 36. 2%).

The results indicate that, due to the larger magnitude of mispricing of high priced stocks, the difference between the return of low priced stocks(overpriced & underpriced) and that of the high priced stocks(overpriced & underpriced) is getting bigger with the increase of holding period. Therefore, this essay used the feature to construct momentum strategies to buy stocks with low magnitude of mispricing and sell stocks with high magnitude of mispricng, that is, buying low priced stocks(overpriced & underpriced) and selling high priced stocks(overpriced & underpriced), generating a total of four sets of momentum strategy. Using a one-month formation period and a one-to-twelve month holding period, we observed if these four strategies were able to yield a significant return. In view of variance of overlapping returns may have bias, in this paper, we used the method by Bod et al. (2002) to correct the variance before looking up the t value.

Table 2 illustrates four momentum returns increase with an increase in the holding period and the average returns are significantly larger than zero. The aggregate amount gained over a 12-month holding period reaches as high as 36.2% when low-priced/underpriced clusters are purchased and high-priced/underpriced clusters are sold. When low-priced/overpriced clusters are purchased and the high-priced/overpriced clusters are sold, there is almost the lowest cumulative return, but it still reaches as high as 20%. This indicates that using mispricing effects to construct a momentum portfolio is an effective strategy.

Table 2. Momentum Strategy-- long Low- price and Short High- price

Holding	Momentum Strategy			
Period	Lo - Hu	Lu - Hu	Lo-Ho	Lu-Ho
1	0.004	-0.017	-0.022	-0.044
2	0.028	0.026	0.006	0.004
3	0.040	0.057	0.011	0.027
4	0.082	0.092	0.047	0.056
5	0.111	0.119	0.070	0.079
6	0.133	0.145	0.096	0.108
7	0.151	0.172	0.111	0.132
8	0.163	0.202	0.132	0.170
9	0.193	0.243	0.145	0.194
10	0.224	0.283	0.161	0.220
11	0.252	0.322	0.182	0.252
12	0.275	0.362	0.200	0.287
Average	0.138	0.167	0.095	0.124
Std.	0.057	0.087	0.041	0.072
t-value	5.382***	4.834***	4.494***	4.137***

Taking factors such as firm size, January effect, multiple bear market and industry into consideration, the paper continued to examine whether market status and corporate characteristics would have an impact on the momentum strategy. First, the companies selected were arranged in order of decreasing market value. Companies in the top 30% were categorized as "big" companies and those in the bottom 30% were classified into "small" businesses. In accordance with the foregoing method, big companies were further divided into high-priced/overpriced cluster, high-priced/underpriced cluster, low-priced/overpriced cluster, and low-priced/underpriced cluster, and four momentum strategies were constructed. Same procedure was applied to small companies and four momentum strategies were also constructed. Secondly, survey of the literature indicated that window dressing by institutional investors and tax-loss selling also contribute to stock return momentum(Haug and Hirschey, 2006；Sias, 2007). To observe the impact of January effect on the moment strategy, we divided the formation period into two types. One formation period includes the month of January and the other does not. Third, the framework proposed by Pagan and Sossounov (2003) was used to differentiate bull and bear markets in Taiwan stock market. Last, according to Moskowitz and Grinblatt (1999), the profits of momentum strategies are known to be correlated with industry. Of the total number of industrial companies, 57% are electronic companies. Electronic industry must have some type of influence on the market. This study sorted all stocks into two groups - electronic stocks and non-electronic stocks - to examine the influence of industry on momentum strategies.

1. Firm size. The empirical results that all four momentum strategies realize extraordinary returns. In the category of small businesses, buying low-priced/overpriced clusters and selling high-priced/underpriced clusters or buying low-priced/underpriced clusters and selling high-priced/underpriced clusters lead to higher cumulative returns. In the category of big companies, buying low-priced/overpriced clusters and selling high-priced/overpriced clusters or buying low-priced/underpriced clusters and selling high-priced/overpriced clusters generate higher cumulative returns. In other words, in the category of small businesses, the momentum profit is driven by high-priced/underpriced clusters, while the high-priced/overpriced cluster is the driver in the category of big companies. As high priced stocks have a larger magnitude of mispricing, the results of the study once again prove that momentum strategy is correlated with mispricing.

2. January effect. There is not much difference of returns between the formation period including the month of January and that without the month of January. Four momentum strategies yield extraordinary returns.

3. Bull market and bear market. During a bear market, four momentum strategies continue to yield extraordinary returns. However, in a bull market, high-priced/overpriced stocks are not as statistically significant as high-priced/underpriced ones.

4. Industry. All momentum strategies for electronic stocks yield extraordinary returns. However, there is something deserving special attention. With non-electronic sample, buying low-priced/overpriced clusters and selling high-priced/underpriced clusters has suffered significant negative returns instead.

CONCLUSIONS

Many of the previous studies on momentum strategies have concentrated primarily on the rollover approach in Jegadeesh and Titman (1993) , that is, buying winner stocks and selling loser stocks, to construct momentum strategies. This study constructed another type of momentum strategy from the perspective of mispricing in company’s stocks. The high magnitude of mispricing often means investors have a greater tendency to overreact, then based on the past mispricing, buying stocks with low magnitude of mispricing and selling stocks with high magnitude of stocks should yield extraordinary returns. Empirical results show that low priced stocks yield higher returns because their misrpricing phenomenon is significantly lower than that of high priced stocks. The implementation of momentum strategies show returns of four momentum strategies increase with an increase in the holding period, indicating that constructing a momentum portfolio based on the magnitude of mispricing is an effective strategy.

Meanwhile, factors such as firm size, January effect, bear market and industry do not have a significant impact except on non-electronic stocks. With non-electronic stocks, buying low-priced/overpriced clusters and selling high-priced/underpriced clusters has a significant negative returns instead.

REFERENCES

Blitzer, D. M. and Dash. S., *Low Price Stocks: Key Characteristics and Indexing Perspective*, Standard & Poor's, 2002.

Bod, P., Blitz, D., and Franses, P. H. "An Unbiased Variance Estimator for Overlapping Returns," *Applied Financial Economics*, 2002, 12(1), 155-158.

Chen, L., Petkova R., and Zhang, L. "The Expected Value Premium," *Journal of Financial Economics*, 2008, 87(2), 269-280.

Frankel, R., and Lee, C. M. "Accounting Valuation, Market Expectation and Cross -sectional Stock Returns", *Journal of Accounting and Economics*, 1998, 25(3), 283-319.

Jegadeesh, N., and Titman, S. "Returns to Buying Winners and Selling Losers: Implication for Market Efficiency," *Journal of Finance*, 1993, 48(1), 65-91.

Jenter, D. "Market Timing and Managerial Portfolio Decisions", *Journal of Finance,* 2005, 60(4), 1903-1949.

Haug, M., and Hirschey, M. "The January Effect," *Financial Analysts Journal*, 2006, 62(5), 78-87.

Kahle, M. K. "Insider Trading and The Long Run Performance of New Security Issues," *The Journal of Corporate Finance*, 2000, 6(1), 25-53.

Lakonishok, J., Shleifer A., and Vishny, R. "Contrarian investment, extrapolation, and Risk," *Journal of Finance*, 1994, 49(5), 1541–1578.

Lakonishok, J., and Lee. I. "Are Insider Trades Informative?" *Review of Financial Studies*, 2001, 14(1), 79-111.

MacKinlay, A. C., and Ramaswamy, K. "Index-Futures Arbitrage and the Behavior of Stock Index Futures Prices," *Review of Financial Studies*, 1988, 1(2), 137-158.

Moskowitz, T. J., and Grinblatt, M. "Do Industries Explain Momentum," *Journal of Finance*, 1999, 54(4), 1249-1290.

Pagan, A. R., and Sossounov, K. A. "A Simple Framework for Analyzing Bull and Bear Markets," *Journal of Applied Econometrics*, 2003, 18(1), 23-46.

Rozeff, M. S., and Zaman, M. A. "Overreaction and Insider Trading: Evidence from Growth and Value Portfolios," *The Journal of Finance*, 1998, 53(2), 701-716.

Sias, R. "Causes and Seasonality of Momentum Profits," *Financial Analysts Journal,* 2007, 63(2), 48-54

Skinner, D. J., and Sloan, R. G. "Earnings Surprises, Growth Expectations and Stock Returns or Don't Let an Earnings Torpedo Sink Your Portfolio," *Review of Accounting Studies*, 2002, 7(2), 289–312.

THE IMPACT OF FINANCIAL CRISIS ON THE VALUE-RELEVANCE OF FINANCIAL INSTRUMENTS FAIR VALUE INFORMATION

Shih-Yung Fen, Department of Accounting, National Yunlin University of Science and Technology, Phone: 886-5-5321612, e-mail:g9620805@yuntech.edu.tw
Shaw K. Chen, College of Business Administration, University of Rhode Island, Kingston,RI02881, Phone: 401-874-4339, e-mail:chenshaw@uri.edu
Chung-Jen Fu, Department of Accounting, National Yunlin University of Science and Technology, Phone: 886-5-5342601#5341, e-mail:drfucj@gmail.com

ABSTRACT

This study investigates how the incremental explanatory power of financial instruments fair value information is affected by the financial meltdown. In addition, the variation in value-relevance of earning and book value of equity are also studied. We extend the empirical model of Ohlson (1995) and our results suggest that during the financial meltdown, the value-relevance of most fair value accounting information on financial instruments decreases significantly. Our result shows that the value-relevance of most accounting information is affected by the major changes in the economic and financial condition. The variation in value-relevance of accounting information is also affected by the characteristics of each industry.

Key word: Value-relevance, Financial instruments, Fair value

I. INTRODUCTION

On September 15, 2008, the century-old Lehman Brothers filed for bankruptcy protection; Merrill Lynch agreed to be merged by Bank of America at US$ 50 billion dollars, and AIG failed to get a US$ 40 billion bailout. The Dow Jones industrial average dropped 504 points that day, and the global stock markets plummeted. The financial meltdown resulting from the U.S. subprime mortgage crisis ultimately grew into a major disaster. The financial crisis in U.S. soon spread to Europe and Asia, and other countries could not escape from this crisis. However, the economic recession and decreased purchasing power in western countries damaged Asian export performance. The shortfall of export orders, coupled with capital withdrawal, caused a sharp decline in Asian stock markets. TAIEX of Taiwan Stock Exchange dropped from 6,310 points on September 12, 2008 to 4,242 points on January 20, 2009, representing a 32.77% decrease.

Accounting standards have moved towards fair value in recent years. Since 1990, the United States has issued SFAS 105, SFAS 107, SFAS 119, SFAS 133, SFAS 157, and SFAS 159 to regulate accounting standards for financial instruments and relevant transactions. IASB issued IAS 32 in 1995 and then issued IAS 39 in 1998. Based on IAS 39, Taiwan issued the Statement of Financial Accounting Standards (SFAS) NO.34 on December 25, 2003, which took effect in 2006. According to SFAS No. 34, financial instruments shall be evaluated at fair value; in addition, financial derivatives shall also be recognized in the statements and evaluated at fair value.

This study focuses on how the value relevance of information of high-tech and low-tech industry (traditional industry) in Taiwan was affected by the financial crisis. According to Lev and Zarowin (1999), when there are material changes to the external environment, investors will immediately evaluate potential impacts on the company and whether or not they will be reflected in the stock price. Since there is a gap between the event and the disclosure of financial statements, the latter cannot immediately reveal external changes. As a result, its value relevance decreases. The financial crisis in 2008 is related to financial instruments. Therefore, we focus on how the relevance of fair value information on financial instruments was affected by the financial crisis.

Our evaluation model is based on Ohlson's model. We integrate the fair value accounting information of financial instruments into our model. With the issuing of fair value accounting standards, the value relevance of fair value accounting information has been studied by many researchers, who mostly focus on the financial industry. Ahmed, Kilic, and Lobo (2006) investigate the value- relevance of banks' derivative financial instruments held for risk management purposes that are recognized in financial statements at fair value and those that are not recognized, but whose fair value are disclosed in notes to the financial statements. Their findings suggest that SFAS No.133 has been successful in increasing the transparency and visibility of derivative financial instruments. The results suggest that recognition and disclosure are not substitutes.

The results of this study show that, the value-relevance of the financial instruments fair value information and earning and book value of equity are affected by the major changes in the economic and financial environment. The variation in value-relevance of accounting information is also affected by the characteristics of each industry.

The rest of this paper is organized as follows. Section 2 presents the development of testable hypothesis. Section 3 describes the research design. Section 4 presents the empirical analysis. Section 5 concludes the paper.

II. DEVELOPMENT OF TESTABLE HYPOTHESIS

First, we study whether or not the fair value accounting information on financial instruments for high-tech and low-tech companies are value-relevant. Then, we analyze whether during the financial crisis, such relevance decreased. Since the fair value accounting reduces the gap between book value and market prices, a company's book value is supposed to be closer to its real economic value. We infer that when economic conditions are relatively stable, the fair value accounting information on financial instruments will be relevant. The following is our hypothesis:

H1: Prior to the financial crisis, the fair-value accounting information on financial instrument was value relevant.

The financial Meltdown in 2008 weakened the international economy, and the international financial market almost collapsed. Under these circumstances, investors may have exaggerated their losses on financial instruments and neglected the hedging effect; accordingly, they may change the stock pricing model and ignored the value of accounting information. Consequently, we assume that the link between stock price and accounting information was affected. We infer that, during the financial crisis, the value relevance of accounting information on financial instruments may decrease. Therefore, we proposed the following hypothesis:

H2: During the financial crisis, the value relevance of accounting information on financial instruments may decrease

Based on discussion above, we speculate that investors to panic and depend on accounting information less when financial and economic conditions are worse. However, the stock pricing model may be adjusted to change the relationship between stock prices and accounting information. Therefore, we proposed the following hypothesis:

H3: During the financial crisis, the earnings and book value of equity were less value-relevant

III. RESEARCH DESIGN

In this study, the fair value accounting information of financial instruments refers to the valuation gains or losses on financial instruments and unrealized gains or losses on financial instruments. These two are the main test variables in this study; however, we are also interested in the earnings and book value. Our model is based on Ohlson (1995), refer to the equity evaluation model in Collins, Maydew, and Weiss (1997). The market value of equity is a function of the earnings and book value of equity. These earnings include the valuation gain (loss) on financial instruments and residual earnings; the book value includes unrealized gain (loss) on financial instruments and residual book value. The following section presents the empirical models used in this study.

$$MV_{it} = \beta_0 + \beta_1 VI_{it} + \beta_2 OE_{it} + \beta_3 UVI_{it} + \beta_4 OBV_{it} + \beta_5 SIZE_{it} + \beta_6 CF_{it} + \beta_7 D1 + \beta_8 D1 * VI_{it} + \beta_9 D1 * OE_{it} + \beta_{10} D1 * UVI_{it} + \beta_{11} D1 * OBV_{it} + \beta_{12} D1 * SIZE_{it} + \beta_{13} D1 * CF_{it} + \varepsilon_{it} \quad (1)$$

In the model, *i* and *t* represent the company and period, respectively. The dependent variable market value (*MV*) is defined as stock price times the number of shares outstanding for firm *i* at the end of period *t*. As for independent variables, the gain or loss is one of the most concerned items of the fair value information on financial instruments. Therefore, we select the valuation gain (loss) on financial instruments and the unrealized gain (loss) on financial instruments as our test variables.[1] We first study the value relevance of these two variables, and then examine whether or not the value relevance decreases during a financial crisis. *VI* denotes the valuation gain (loss) on financial instruments for period *t*. *OE* is pre-tax income minus *VI*. In previous studies, some researchers used the Ohlson equity evaluation model for empirical analysis, replacing abnormal earnings with current earnings (Aboody 1996; Barth et al., 1998; Chen, 2003). We use quarter earnings to replace abnormal earnings in this study. *UVI* is the unrealized gain (loss) on financial instruments. *OBV* is the book value minus *UVI*. In addition, we use the control variables, *SIZE* and *CF*, to control potential price impacts resulting from the company size and cash flow. *SIZE* is total assets at period-end, and *CF* is the operating cash flow. All variables are deflated by book value of equity at the beginning of period *t*. We use the book value of equity to be the deflator to control the size or scale effect and mitigate the heteroscedasticity resulting from different company sizes. To test how the financial crisis impacts the value relevance of accounting information, we add a dummy variable, *D1*, to the model. *D1* = 0 before the crisis, and *D1=1* during the crisis. We expect $\beta_7, \beta_8, \beta_9, \beta_{10}, \beta_{11}$ to be significantly negative in model 1, representing a decrease in the value relevance of the following accounting information during the financial crisis: fair value accounting on financial instruments, earnings, and book value. And, we do not predict the

[1] Valuation gain(loss) on financial instruments include Valuation gain on financial Assets 、Valuation gain on financial Liabilities 、Valuation loss on financial Assets and Valuation loss on financial Liabilities. Unrealized gain(loss) on financial instruments include Unrealized gain on financial Assets 、Unrealized gain on financial Liabilities 、Unrealized loss on financial Assets and Unrealized loss on financial Liabilities.

signs of the *SIZE* and *CF* estimated coefficients.

The study period starts from the first quarter of 2007 to the first quarter of 2009. The total sampling period covers 9 quarters, of which 6 (Q1, 2007 ~ Q2, 2008) are not during the financial crisis, and 3 (Q2, 2008 ~ Q1, 2009) are during the crisis. Our samples are TSEC- and OTC-listed high-tech and low-tech companies in Taiwan. The data in this study comes from the Taiwan TEJ equity and Taiwan TEJ Finance; we also cross-reference data from the TWSE database.[2] Extreme values were excluded, and we have a total of 6,358 firm-quarter observations for high-tech industries, and 4,102 firm-quarter observations for low-tech industries.[3]

IV. EMPIRICAL ANALYSIS

We used the ordinary least squares (OLS) model in our multiple regression analysis. By using the econometric software EViews, the coefficients of the regression model are estimated as shown in Table 1. Table 1 presents the regression results of high-tech and low-tech samples before the financial crisis. According to Panel A (high-tech sample), the coefficient of *VI* and *UVI* are significantly larger than 0, suggesting that both *VI* and *UVI* of high-tech companies are useful in explaining firm values. According to Panel B (low-tech sample), the coefficient of *VI* is positive but not statistically significant. *UVI* are significantly larger than 0, suggesting *UVI* of low-tech companies is also useful in explaining firm values.

Two items in the fair value information on financial instruments for the high-tech industry and the *UVI* for the low-tech industry have incremental explanatory powers on stock prices, but not the *VI* in low-tech industry. A potential reason is that the amount of *VI* is usually relatively small. Compared to the high-tech companies, the low-tech companies have fewer financial assets and financial liabilities that create *VI*. As compared to the high-tech industry, the low-tech industry is often domestic and has fewer imports and exports; thus, companies do not often engage in currency hedging. In addition, the low-tech companies are not as active in research and development as compared to the high tech companies, so the former does not have frequent capital needs and issues fewer convertible bonds. Financial assets and liabilities occur when a company uses currency hedging or issues convertible bonds, creating valuation gains or losses under fair value accounting standards.

The regression results in Panel A of Table 1 show that during the financial crisis, the coefficient of *D1*VI* is significantly smaller than 0 at a.01 level, and supporting H2. The coefficient of *D1*UVI* is positive, which is not consistent with the expectations. Thus, H2 was supported partially in high-tech sample. According to Panel B, the coefficient of *D1*VI* is negative, and marginally significant. The coefficient of *D1*UVI* is significantly smaller than 0 at a.01 level. Thus, findings in low-tech sample, H2 received fully support. Even though the financial crisis has affected the high-tech and traditional industries differently, with regard to the value relevance of fair value information on financial instruments, the relevance of *VI* was lost during the crisis in two tests, while *UVI* still remained relevant.

On the other hand, the coefficient of *D1*OE* in Panel A is significantly smaller than 0, and

---

[2] TWSE is Taiwan stock exchange.

[3] High-Technology samples include Biotechnology and Medical Care, Electricity Industry, Semiconductor Industry, Computer and Peripheral Equipment, Optoelectronic Industry, Communications and Internet, Electronic Parts/Components, Electronic Products Distribution, Information Service Industry and Other Electronic Industry. Low-Technology samples include Cement, Food, Plastic, Textile, Chemical Industry, Electric Machinery, Electric & Cable, Glass & Ceramic, Paper & Pulp, Iron & Steel, Rubber, Automobile, Building Material & Construction, Shipping & Transportation, Tourism, Trading & Consumers Goods, General and Other Industry.

H3 was supported partially in high-tech sample. The coefficients of *D1*OE* and *D1*OBV* in Panel B are both significantly smaller than 0, and H3 was supported fully in low-tech sample. Furthermore, the results in Table 2 indicate that most of the value relevance of earnings and book value reduce as a result of the crisis. In summary, our results indicate that the relevance of fair value information on financial instruments of low-tech companies was affected by the financial crisis significantly, which is consistent with the expectations.

V. **CONCLUSION**

This study examines whether or not the fair value accounting information for the high-tech and low-tech industries, after the implementation of SFAS No. 34 in Taiwan in 2006, became less value-relevant as a result of the global financial Meltdown in 2008. In conclusion, most fair value accounting information on financial instruments is value-relevant when the economy is stable. The relevance of most fair value information on financial instruments is affected by the financial crisis, so is the value relevance of earnings and book value. As a result, the explanatory power of accounting information on stock price is lower. Furthermore, the variation in value-relevance of accounting information is also affected by the characteristics of each industry.

Accounting information becomes less important to investors when economic or financial conditions change dramatically; and fair value accounting information on financial instruments becomes irrelevant and has no incremental explanatory power on the stock price. The results of this study confirm the hypotheses, showing that the accounting standards meant to make information more effective can still be improved. During the Financial Meltdown, IAS 39 and SFAS No. 34 of Taiwan were amended in October 2008 to alleviate the impacts of fair value accounting. However, the value relevance of accounting information was not significantly improved, and the amendments were not very effective in restoring investor confidence. In the future, as economic conditions become more complex and the world economy becomes more integrated, a financial crisis in a single country is likely to spread to other countries, creating a global disaster. The financial Meltdown in 2008 was an obvious example. In the changing economy, it is important to study how accounting information responds to environmental changes and how to improve the information value and avoid price huge fluctuation as a result of fair value accounting.

REFERENCE

［1］Ahmed, A.S., E. Kilic, and G.J.Lobo. “Does recognition versus disclosure matter? Evidence from value-relevance of bank’ recognized and disclosed derivative financial instruments,” *The Accounting Review* ,2006, 81,567-588.

［2］Barth M. E., W. H. Beaver, and W. R. Landsman. “Value-Relevance of Banks’ Fair Value Disclosures under SFAS No.107,” *The Accounting Review*, 1996, 71,513-537

［3］Collins, D., C. Maydew, and I. Weiss.. “Changes in the value-relevance of earnings and book values over the past forty years,” *Journal of Accounting and Economics*, 1997, 24, 39-67.

［4］Core,J, W. Guay and A. Buskirk.. “Market Valuation in the New Economy: An Investigation of What has Changed,” *Journal of Accounting and Economics*, 2003, 34, 43–67.

［5］Francis,J., and K. Schipper.. “Have financial statements lost their relevance？”*Journal of Accounting Research*, 1999, *37 (Supplement)*: 319-352.

［6］Lev, B., and P. Zarowin. “The boundaries of financial reporting and how to extend them,” *Journal of Accounting Research*,1999, 37 (2): 353-389

［7］Wang,L. P. Alam, and S.Makar.. “The value-relevance of derivative disclosures by commercial banks: A comprehensive study of information content under SFAS Nos.119 and 133,” *Review of Quantitative Finance and Accounting* , 2005, 25(4):413-427

TABLE 1

OLS regression-primary test of hypothesis 1 、2 and 3

Value-relevance of Accounting Information for the Pre- and During financial Meltdown

		Panel A--Hi-tech industry				Panel B--Low-tech industry			
Variable	Predicted Sign	Coefficient	t-Statistic		p-value	Coefficient	t-Statistic		p-value
$1/BV$		149458.0	6.914	***	0.000	109996.8	4.475	***	0.000
VI_{it}	+	5.648	1.889	*	0.059	1.447	0.628		0.530
OE_{it}	+	1.227	3.033	***	0.002	3.868	4.115	***	0.000
UVI_{it}	+	2.872	5.327	***	0.000	1.961	11.207	***	0.000
OBV_{it}	+	2.282	36.026	***	0.000	1.227	18.752	***	0.000
$SIZE_{it}$	?	-0.087	-2.574	***	0.010	0.077	2.442	**	0.015
CF_{it}	?	0.695	3.995	***	0.000	-0.136	-0.582		0.561
$D1$	—	-3.079	-5.235	***	0.000	0.743	1.503		0.133
$D1*VI_{it}$	—	-13.412	-2.933	***	0.003	-5.093	-1.642	*	0.101
$D1*OE_{it}$	—	-5.204	-4.875	***	0.000	-2.163	-2.092	**	0.037
$D1*UVI_{it}$	—	0.843	1.045		0.296	-1.623	-3.474	***	0.001
$D1*OBV_{it}$	—	2.007	2.982	**	0.003	-1.063	-2.165	**	0.030
$D1*SIZE_{it}$	?	-0.098	-1.213		0.225	-0.099	-2.984	***	0.003
$D1*CF_{it}$	?	0.400	0.868		0.385	0.141	0.486		0.627
Adjusted R-squared		0.39				0.27			
Number of observations		6358				4102			

* (**,***) significant at 10% (5%,1%) level (two-tailed)

All p-values are based on White (1980) heteroscedastcity corrected t-value.

i,t = firms and quarters, respectively

MV_{it} = market value of common equity at the end of quarter t / book value at the beginning of quarter t

VI_{it} = Valuation Gain or Loss on financial instruments in the quarter t / book value at the beginning of quarter t

OE_{it} = (pre-tax net income in the quarter $t - VI$) / book value at the beginning of quarter t

UVI_{it} = Unrealized Gain or Loss on financial instruments at the end of quarter t / book value at the beginning of quarter t

OBV_{it} = (book value at the end of the quarter t- UVI) / book value at the beginning of quarter t

$SIZE_{it}$ = total assets at the end of quarter t / book value at the beginning of quarter t

CF_{it} = cash flow from operating activities in the quarter t / book value at the beginning of quarter t

$D1$ = dummy variable D1=1 if during the financial crisis , and 0 otherwise.

THE IMPACT OF LIFE CYCLE ON THE VALUE RELEVANCE OF THE KEY COMPONENTS OF FINANCIAL STATEMENT

Shaw K. Chen, College of Business Administration, URI, (401) 874-4339, chenshaw@uri.edu
Yu-Lin Chang, Department of Accounting and Information Technology, Ling Tung University, (886) 4-23892088, g9320821@yuntech.edu.tw
Chung-Jen Fu, Department of Accounting, National Yunlin University of Science and Technology, (886) 5-5342601#5511, drfucj@gmail.com

ABSTRACT

The components of earnings or cash flows have different implications for the assessment of firm value. We extends the search for value-relevant fundamentals to examine which financial performance measures can convey more information to help investors evaluate the performance and value for firms of different life cycle stages in the high-tech industry. Six performance measures are used to explain the difference between market value and book value, and cross-sectional data from firms in Taiwanese information electronics industry are used to test our hypotheses. We find all the six performance measures which are decomposed from Income Statement and Cash Flow Statement are important value indicators but the relative degrees of value relevance of various performance measures are different across the firm's lifecycle stages. The market does concern with the various financial performance measures in different lifecycle stages and reflects it on the stock price.

1. INTRODUCTION

Any variable is considered value relevant if it has a predictable association with market values of equity. A sizable literature suggests that financial measures can provide value-relevant information for investors. The components of earnings or cash flows have different implications for the assessment of firm value. Firms in different life cycle stages have different economic characteristics that may affect the usefulness and value relevance of financial performance measures (Black, 1998; Baginski, Lorek & Branson, 1999). It is necessary to consider the impact of the life cycle stage on the value relevance of the components of earnings and cash flows.

This study uses cross-sectional data over twelve years to examine six financial performance measures for firms of different life cycle stages in the Taiwanese high-tech industry. The earning measures are decomposed into research and development expense (R&D), operating income (OI) and adjusted non-operating income (ANOI). The cash flow measures are decomposed into cash flows from operating (CFO), cash flows from investing (CFI) and cash flows from financing (CFF).

We find all the six performance measures are important value indicators but the relative degrees of value relevance are different across firm lifecycle stages. Our result supports the Financial Accounting Standard Board (FASB) suggestion that present and potential investors rely on accounting information to improve investment, credit, and similar predictive decisions. Specifically, our findings detail the value relevance of different financial performance measures across lifecycle stages of information electronic firms, a critical and highly competitive industry for Taiwan's economy.

The remainder of this paper is organized as follows. Section II discusses the literature review and develops hypotheses from past research, Section III presents our sample selection procedure, lifecycle classification, definition and measurement of the variables and the empirical model. The empirical results and analyses are discussed in section IV. Further discussion on our finding and a brief conclusion are provided in the final section.

2. LITERATURE REVIEW AND HYPOTHESES

Financial performance measures have been most widely used to evaluate the organizational health of a firm. We expect all these six key financial measures are positively associated with firm value in different life cycle stages. However, the opportunities, pressures, and threats in both the external and internal environment of an organization vary with the stages of life cycle (Anderson & Zeithaml, 1984; Jawahar & Mclaughlin, 2001). The information conveyed by the financial performance measures is expected to be different for each component in different life cycle stages. We develop related hypotheses across the different lifecycle stages.

2.1 The Growth Stage

R&D is a critical element in the production function of information electronic firms. Booth (1998) and Hand (2003, 2005) finds that a firm with continuous research and development will have a higher ratio of hidden value in its firm value and supports the proposition that R&D is beneficial to the future development of the firm. Relative to the later two lifecycle stages, firms in the growth stage will spend more on R&D leading to future profit-generating opportunities and increase firm value. This discussion leads to the following hypothesis:

H_{1a}: For firms in the growth stage, R&D is more positively associated with firm value.

In the growth stage, most firms can't generate much income from core operating activities. Those who can obtain more income from non-operating activities will create more value shareholders. Fairfield et al. (1996) suggest that non-operating income has incremental predictive content of future profitability. Therefore, in this stage, the firm's adjusted non-operating income (ANOI) is positively associated with firm value and we propose the following hypothesis:

H_{1b}: For firms in the growth stage, ANOI is more positively associated with firm value.

2.2 The Mature Stage

Robinson (1998) indicates that operating income (OI) is a superior measure to reflect a firm's ability to sell its products and to evaluate operational performance. Relative to the growth and decline stages, OI is an important earnings item and firms in the mature stage would alter its focus from increasing sales growth to increasing OI. As OI increases, the value of the firm will increase. Thus we test the following hypothesis:

H_{2a}: For firms in the mature stage, OI is more positively associated with firm value.

Klein and Marquardt (2006) views CFO as a financial measure of the firm's real performance. In the mature stage, firms are often characterized by generating sufficient cash flows, without particularly attractive investment opportunities (Jawahar & Mclaughlin, 2001; Penman 2001). Black (1998) also finds that CFO is positively associated with market value of firm in the mature stage. Therefore, as CFO increases, the value of the firm will increase, and leads to our next hypothesis:

H_{2b}: For firms in the mature stage, CFO is more positively associated with firm value.

2.3 The Decline Stage

Business risk is high when firms move into the decline stage. When demand for an organization's traditional products and services are reduced, less efficient firms are forced out of industries and market (Konrath, 1990; Pashley & Philippatos, 1990). Following the BCG (Boston Consulting Group) model, when companies step into the decline stage, the cash generate ability becomes a more important value driver for those companies or divisions in the decline stages. However, declining firms do not necessarily fail. Some firms are forced to consider investment and develop new products and technology to ensure organizational survival (Kazanjian, 1988; Jawahar & Mclaughlin, 2001; Zoltners, Sinha & Lorimer, 2006). Black (1998) points out that firms can regenerate by investing in new production facilities and innovative technology and goes back into the growth or mature stage, prevent failure for many years. However, if the non-productive and idle assets are disposed and managers make decisions in the investor's best interest, then CFI is expected to be positively correlated with firm value during decline (Black, 1998). Therefore, regardless of whether the firm's CFI increases or decreases, as managers make appropriate decisions strategically, the value of the firm will increase. Thus, we develop the following hypothesis:

H_{3a}: For firms in the decline stage, CFF is more positively associated with firm value.

H_{3b}: For firms in the decline stage, CFI is more positively associated with firm value.

3. RESEARCH DESIGN AND METHODOLOGY

3.1 Sample Selection and Lifecycle Classification

We select publicly-listed information electronics companies from the Taiwan Stock Exchange (TSE) and Gre Tai Securities Market (Taiwan OTC market). The companies' financial data and the equity market value data are obtained from the Financial Data of Company Profile of the Taiwan Economic Journal (TEJ) Data Bank. A total of 4,862 firm-year observations over twelve years from 1997 to 2008 are collected.

Classifying companies into different life cycle stages is a challenging task**.** This study uses a tailored classification method similar to Anthony and Ramesh (1992) and Black (1998). A multivariate classification method is used to classify observations into three life cycle stages[1]. Our final sample retains 1,070 firm-year observations in growth stage, 2,730 firm-year observations in mature stage and 1,062 firm-year observations in decline stage.

3.2 Empirical Model

We extend Ohlson's(1995) valuation model to examine the relationship between equity market value and the various financial performance measures in each life cycle stage. We add a dummy variable (Stage) based on the lifecycle stage; 1 for the earlier life cycle stage and 0 for the latter life cycle stage, and build two-way interaction terms among various financial variables to test our hypotheses. Definition and measurement of variables are given in Table 1. The extended empirical model is as follows.

$$MV_i = b_0 + b_1BVC_i + b_2CFO_i + b_3CFO_i{*}Stage_i + b_4CFI_i + b_5CFI_i{*}Stage_i + b_6CFF_i + b_7CFF_i{*}Stage_i + b_8R\&D_i + b_9R\&D_i{*}Stage_i + b_{10}OI_i + b_{11}OI_i{*}Stage_i + b_{12}ANOI_i + b_{13}ANOI_i{*}Stage_i + \varepsilon_i$$

(**Insert Table 1 about Here)**

4. ANALYSIS OF EMPIRICAL RESULTS

We examine the incremental effect of different life cycle stages on the financial performance measures and show the compared results between each two life cycle stages in Table 2. The empirical results show the estimated coefficient of the six financial measures are all significant positively associated with firm value in each stage alone.

(**Insert Table 2 about Here)**

The estimated coefficient of the interaction term R&D*Stage is significant (Panel A and t = 8.645, p<0.01; Panel C and t = -6.901, p<0.01). This result supports our inference

[1]The procedure of lifecycle classification is as follows. First, we choose sales growth, capital expenditures, dividend payout, and firm age as the classification indicators. Second, sales growth and capital expenditure are sorted from highest to lowest, while dividend payout and firm age are sorted from lowest to highest by rank. The indicators are given a score of 0, 1 or 2 based on their ranking. The firm with the highest sales growth or capital expenditures is given a score 0; or else, the firm with the highest dividend payout or firm age is given a score 2. The scores of the four classification indicators are then summed together giving a composite score that ranges from zero to eight. Finally, firm-years are assigned to one of the three groups based on the composite score. Firm-year observations with composite score ≦ 2 are assigned to the growth stage. Firm-years with a composite score three, four, or five are assigned to the mature stage. And, firm-year observations with composite score ≧ 6 are assigned to the decline stage.

that R&D carries more weight for the growth stage than the other two lifecycle stages. It also implies the R&D investments, especially for the firms in the growth stage, are critical for firm valuation. In addition, the estimated coefficient of the interaction term ANOI*Stage is also significant (Panel A and t = 7.677, p<0.01; Panel C and t = -6.650, p<0.01). These results indicate that the market has a higher positive valuation for firms which have higher R&D and ANOI in the growth stage. The results support H_{1a} and H_{1b}.

The estimated coefficient of OI*Stage in the mature is more significant relative to the decline stage (Panel B and t = 4.806, p<0.01). This result reflects a fact for Taiwanese information electronics firms in the mature stage that with the ability to generate higher operating income have more positive valuation than firms in the decline stage. But the estimated coefficient of OI*Stage in Panel A is positive which means the growth stages are more significant relative to the mature stage. The results only partially support H_{2a}. On the other hand, the estimated coefficient of CFO*Stage in the mature stage is higher than the other two stages but not significant (Panel A and t = -1,206, p>0.1; Panel B and t = 0.188, p>0.1). The result is not support H_{2b}.

Comparing CFF*Stage among the three stages, we find the estimated coefficient of CFF*Stage in the decline is more value relevant than in the growth stage and mature stage (Panel B and t = -0.101, p>0.1; Panel C and t = 2.626, p<0.01). The results partially support H_{3a}. The estimated coefficient of CFI*Stage in the decline stage is positively negatively significant than in the growth stage (Panel C and t = 1.715, p<0.1), but is smaller than in the mature stage. The results partially support H_{3b}. In sum, we find that the impact of lifecycle exists and is helpful to explain the association between various performance measures and firm value among Taiwanese firms.

5. CONCLUSION

This study examines the financial performance measures in different life cycle stages to identify the key measures for investors when they value the firm using the data from firms in Taiwan's information electronic industry. We find that R&D, ANOI, CFI and CFF are value-relevant in the growth stage; CFO and OI are value-relevant in the mature stage; and CFI can convey information on the firm's value in the decline stage.

Our empirical evidence show a differential role of financial performance measures across life cycle stages, and highlights the role of cash flows and various earnings-related items in explaining firm market value. Our findings differ from Liang and Yao's (2005) arguments that traditional financial statements face an unprecedented challenge of adequately evaluating the high-tech industry's firms' value. Our findings contribute to financial performance measures value-relevant research, which can be of interest to a broad public, including academic researchers, standard setters, financial statement preparers and users, and the practice.

The result in this study is specifically directed to the information electronics industry in Taiwan. As the business strategies of the Taiwanese information electronics industry are multifaceted and different firms have diverse specializations, we see benefits for future research to classify and identify firms based on the product lines and the distribution channel positions in the supply chain.

REFERENCES

Anderson, C. R., and Zeithaml, C. P., 1984, "Stage of the Product Life Cycle, Business Strategy, and Business Performance," Academy of Management Journal 27 (March): 5-24.

Anthony, J. H., and Ramesh, K., 1992, "Association between Accounting Performance Measures and Stock Prices," Journal of Accounting and Economics 15 (August):203-227.

Baginski, S. P., Lorek, K. S., and Branson, B. C., 1999, "The Relationship between Economic Characteristics and Alternative Annual Earnings Persistence Measures," The

Accounting Review 74 (January):105-120.
Black, E. L., 1998, “Life-Cycle Impacts on the Incremental Value-Relevance of Earnings and Cash Flow Measures,” Journal of Financial Statement Analysis, 4(fall): 40-56.
Booth, R., 1998, “The Measurement of Intellectual Capital,” Management Accounting, 76 (November): 26-28.
Fairfield, P. M., Sweeney, R. J., and Yohn, T. L., 1996, “Accounting Classification and the Predictive Content of Earnings,” The Accounting Review 71(July): 337-355.
Hand, J. R. M., 2003, “The Relevance of Financial Statements within and across Private and Public Equity Markets,” Working paper, Kenan-Flagler Business School, UNC Chapel Hill.
_____, 2005, “The Value Relevance of Financial Statements in the Venture Capital Market,” The Accounting Review 80(April): 613-648.
Jawahar, I. M., and Mclaughlin, G. L., 2001, “Toward a Descriptive Stakeholder Theory: An Organizational Life Cycle Approach,” Academy of Management Review 26 (July): 397-414.
Kazanjian, R.K., 1988, “The Relation of Dominant Problems to Stage of Growth in Technology-Based New Ventures,” Academy of Management Journal 31 (June): 257-279.
Klein, A., and Marquardt, C., 2006, “Fundamentals of Accounting Losses,” The Accounting Review 81 (January): 179-206.
Konrath, L., 1990, “Audit Risk Assessment: A Discussion and Illustration of The Interrelated Nature of Statements on Auditing Standards,” The Women CPA (summer): 14-16.
Liang, C.J. and Yao, M.L., 2005, “The Value-Relevance of Financial and Nonfinancial Information-Evidence from Taiwan’s Information Electronics Industry,” Review of Quantitative Finance and Accounting 24(March): 135-157.
Ohlson, J., 1995, “Earnings, Book Values and Dividends in Security Valuation,” Contemporary Accounting Research 11(Spring): 661-687.
Pashley, M. M., and Philippatos, G. C., 1990, “Voluntary Divestitures and Corporate Life Cycle: Some Empirical Evidence, “Applied Economics 22 (September): 1181-1196.
Penman, S.H., 2001, Financial Statement Analysis and Security Valuation, McGraw-Hill Irwin.
Robinson, K.C., 1998, “An Examination of the Influence of Industry Structure on the Eight Alternative Measures of New Venture Performance for High Potential Independent New Ventures,” Journal of Business Venturing 14 (March): 165-187.
Zoltners, A. A., Sinha, P., and Lorimer, S. E., 2006, “Sales Force Structure to your Business Life Cycle,” Harvard Business Review 84 (July-August): 81-89.

Table 1. Definition and Measurement of Variables

Variables	Measurement
A.Life cycle classification indicator variables	
Sales growth (SG_{it})	100 * ($sales_t$- $sales_{t-1}$) / $sales_{t-1}$
Dividend payout (DP_{it})	100 * annual dividend of common stock / annual income
Capital expenditures(CE_{it})	100 * (purchase fixed assets – reevaluated fixed assets of firm i at time t) / AV_{t-1}
Firm age	the difference between the current year and the year which the firm was originally formed
Lifecycle-stage($Stage_{it}$)	Dummy variable in the three groups (growth stage compared with mature stage; mature stage compared with decline stage; decline stage compared with growth stage) and take on the value of 1 if the firm operates in the front stage and 0 the other stage.
B. Variables of the empirical model	

Market value of equity (MV_{it})	the market value of equity of firm i at time t / AV_{t-1}
Cash flows from operating (CFO_{it})	cash flows from operating activities of firm i at time t / AV_{t-1}
Cash flows from investing (CFI_{it})	cash flows from investing activities of firm i at time t / AV_{t-1}
Cash flows from financing (CFF_{it})	cash flows from financing activities of firm i at time t / AV_{t-1}
R&D expense ($R\&D_{it}$)	R&D expense of the firm i at time t / AV_{t-1}
Operating income (OI_{it})	(Gross profit – operating expenses) of firms i at time t / AV_{t-1}
Operating income before discontinued and extraordinary items (NI_{it})	the operating income before discontinued and extraordinary items of firms i at time t / AV_{t-1}
Adjusted non-operating income($ANOI_{it}$)	(NI_{it} - OI_{it} – $R\&D_{it}$) /AV_{t-1}
Control variable	
Book value of net assets except for cash(BVC_{it})	the book value of equity less the change in the cash account of firm i at time t / AV_{t-1} (Black, 1998)
Book value of net assets(BV_{it})	the book value of equity of firm i at time t / AV_{t-1}

a. All of the variables are deflated by each year by the book value of assets at the end of year t-1(AV_{t-1}).
b. Except for MV(market value of equity) is measured in million dollars, All of the other variables (CFO, CFI, CFF, R&D,OI, ANOI, BVC, BV, NI and AV_{t-1}) are measured in thousand dollars.

Table 2.The Empirical Results of Incremental Effect of Variables between Stages

$$MV = b_0 + b_1BVC + b_2 CFO + b_3 CFO*Stage + b_4 CFI + b_5 CFI*Stage + b_6 CFF + b_7 CFF*Stage + b_8R\&D + b_9R\&D*Stage + b_{10} OI + b_{11} OI*Stage + b_{12}ANOI + b_{13} ANOI*Stage$$

Panel A: The Growth stage(Stage=1) compared with The Mature stage (Stage=0)			Panel B: The Mature stage(Stage=1) compared with The Decline stage (Stage=0)		Panel C: The Decline stage(Stage=1) compared with The Growth stage (Stage=0)	
Variable	Expected sign	Coefficient (t-Statistic)	Expected sign	Coefficient (t-Statistic)	Expected sign	Coefficient (t-Statistic)
Intercept	?	-0.159 (-2.682***)	?	-0.076 (-1.581)	?	-0.245 (-2.809***)
BVC	+	1.472 (16.058***)	+	1.262 (16.820***)	+	1.618 (12.108***)
CFO	+	1.998 (7.448***)	+	1.676 (4.341***)	+	1.619 (5.876***)
CFO*Stage	–	-0.443 (-1.206)	+	0.188 (0.440)	?	0.149 (0.243)
CFI	+/–	1.887 (6.753***)	+/–	1.575 (4.139***)	+	0.960 (3.548***)
CFI*Stage	+	-1.031 (-2.933***)	+/–	0.024 (0.056)	?	0.995 (1.715*)
CFF	+	2.581 (10.571***)	+	2.538 (7.916***)	+	1.389 (6.765***)
CFF*Stage	+	-1.210 (-3.888***)	?	-0.101 (-0.277)	?	1.299 (2.626***)
R&D	+	5.474 (8.282***)	+	4.071 (4.078***)	+	14.011 (15.055***)
R&D*Stage	+	8.835 (8.645***)	?	2.177 (2.077**)	–	-11.206 (-6.901***)
OI	+	5.259 (22.111***)	+	3.498 (8.980***)	+	7.609 (23.503***)
OI*Stage	–	2.482 (6.937***)	+	2.003 (4.806***)	?	-4.434 (-7.177***)
ANOI	+	1.972 (5.995***)	+	2.069 (6.423***)	+	6.676 (10.809***)
ANOI*Stage	+	4.915 (7.677***)	–	0.163 (0.419)	?	-4.922 (-6.650***)
adj.R^2		0.590	adj.R^2	0.536	adj.R^2	0.628
F value		422.017	F value	337.498	F value	278.152
Number of observables		3,800	Number of observables	3,792	Number of observables	2,132

[a]: *p <0.10 ; ** p <0.05; *** p <0.01
[b]: MV= market value of equity; BVC =Book value of net assets except for cash; CFO =cash flows from operating; CFI =cash flows from investing; CFF= Cash flows from financing; R&D =R&D expense; OI= Operating income; ANOI= Adjusted non-operating income; CFO*Stage= the interaction term of CFO and Stage; CFI*Stage= the interaction term of CFI and Stage; CFF*Stage= the interaction term of CFF and Stage; R&D*Stage= the interaction term of R&D and Stage; OI*Stage= the interaction term of OI and Stage; ANOI*Stage = the interaction term of ANOI and Stage; Stage= life cycle stage.

The Investor Sentiment Spillover Effects of the Stock Index and Stock Index Futures Markets of Taiwan

Yu-Min Wang, National Chiayi University, Taiwan, +886-5-2732826, yumin@mail.ncyu.edu.tw
Chun-An Li, National Yunlin University of Science and Technology, Taiwan, +886-5-5342601, liica@yuntech.edu.tw
Cha-Fei Lin, National Yunlin University of Science and Technology, Taiwan, +886-5-5342601, g9524803@yuntech.edu.tw

ABSTRACT

This study investigates the sentiment spillover effects in the Taiwan spot and futures markets covering the period from 2 August 1999 to 28 February 2009. We find that in a bullish market, virtually all investor sentiment can raise spot returns, although no significant relationship is found between investor sentiment and futures returns. We further examine the overnight returns to determine whether the spillover effects are of a short- or long-term nature, and find that the spot sentiment index becomes weaker, with reduced explanatory power, indicating that the sentiment spillover effects are indeed short-term. Finally, following the application of the pure mean and pure volatility models, we attempt to determine whether asymmetric sentiment volatility exists within the futures and spot markets. Our findings indicate that sentiment in the financial markets can be explained by both returns and the volatility effect.

1. INTRODUCTION

It is quite clear that there has been a considerable increase in focus on the issue of the relationship between the stock and futures markets over recent decades, particularly in the aftermath of the stock market crash of October 1987, an event which subsequently prompted a wealth of related literature.[1] Many of the related studies also began to focus on the linkages and interactions between stock market

[1] Examples include Aggarwal (1988), Edwards (1988), Damodaran (1990), Lee and Ohk (1992), Bessembinder and Seguin (1992), Antoniou and Holmes (1995), Pericli and Koutmos (1997) and Antoniou et al. (1998).

and futures market returns,[2] whilst specific emphasis has also been placed on the effects of investor sentiment on stock prices.

Nevertheless, the prior studies have largely ignored the spillover effects of investor sentiment between the stock market and the futures market.[3] Futures trading, which provides an interactive method of trading between the stock and futures markets, represents an important channel by which investors can engage in both arbitrage and hedging, and since investor sentiment can be communicated within both of these markets, this can clearly affect both returns and volatility.

Whilst the majority of the prior studies tend to focus on the interactions between the stock and futures markets, the present study differs from many of its forerunners in several respects. Firstly, we place particular focus on overnight returns, which helps us to distinguish between the contemporaneous correlations that exist between the two markets. Secondly, we estimate the spillover effects of investor sentiment within both the futures market and the stock market. Thirdly, we expect to find that the spillover effects of investor sentiment will be quite short-term and that they will have quite a rapid effect on prices, which should help investors to make their important investment decisions. Finally, we argue that there is a requirement for investors to adjust their portfolio insurance within the financial market in order to avoid taking on too much risk.

2. LITERATURE REVIEW

In their attempts to ascertain the relationships between the futures market and the stock market, there has apparently been no related discussion within the extant literature on the transmission effects of investor sentiment across these two distinct markets. Therefore, in our review of the prior literature, our focus is on the interactions between the volatility and investor sentiment effects on the prices of financial assets.

Within the extant literature on the futures and spot markets, the majority of the early studies have tended to search for simultaneous and lead-lag relationships between futures and spot prices, or on their respective returns. Both Chan (1992) and Fleming et al (1996) found that S&P 500 index futures tended to lead the stock market; and indeed, in the aftermath of the major collapse which occurred in the US

[2] See, for example, Kock and Koch (1987), Stoll and Whaley (1990), Chan (1992), Lee and Lin (1994), Crain and Lee (1995), Koutmos and Tucher (1996) and Wang (2001).

[3] Examples include Brown (1999), Fisher and Statman (2000), Lee, Jiang and Indro (2002), Baker and Wurgler (2004), Brown and Cliff (2004), Baker and Stein (2004), Charoenrook (2005) and Kumar and Lee (2006).

stock market in 1987, the key points which arose in the subsequent studies were whether trading in the futures market could actually have resulted in an unstable stock market.[4] However, throughout this string of the research, the discussion has invariably tended to focus on the effects of the futures market on the stock market, with very few studies having attempted to pursue any discussion on the potential interdependence between the two markets.

Within the studies outlined above, the focus has generally been placed on the discussion of the correlation which exists between investor sentiment and the stock market. Using the positions held by large traders as a proxy for investor sentiment in order to examine the correlation between such sentiment and the futures market.

In the majority of the prior studies on the effects of investor sentiment in the financial markets, there has been a tendency focus on the relationship between investor sentiment and the stock market; and indeed, there appear to be no studies within the prior literature on the investor sentiment spillover effects between the spot and futures markets. We therefore take this issue into consideration in the present study, and expect to find support for the spillover effects and influence of investor sentiment on the spot and futures markets.

3. DATA AND METHODOLOGY

The data used in this study are daily closing observations covering the period from 2 August 1999 to 28 February 2009, including the TX, TXE and TXF markets, all of which are obtained from the Taiwan Economics Journal (TEJ) database.

It is suggested in some of the prior studies that effective proxies of sentiment could be utilized as time-series variables (Wang et al., 2000; Simon and Wiggins, 2001). In the present study, we similarly adopt the put-call ratio of TAIEX options as a proxy for investor sentiment, measuring the sentiment proxies of the spot market ($SI_{S,t}$) and futures market ($SI_{F,t}$) as follows:

$$SI_{S1,t} = \frac{\text{Rise in the number of companies listed on the TSE}}{\text{All companies listed on the TSE}} \tag{1}$$

$$SI_{S2,t} = \text{Margin Lending Ratio} = \frac{\text{Margin lending on funds}}{\text{Margin lending on securities}} \tag{2}$$

[4] See Antoniou and Holmes (1995), Pericli and Koutmos (1997) and Antoniou et al. (1998).

$$SI_{S3,t} = \text{Put-Call Ratio} = \frac{\text{Total trading volume of puts}}{\text{Total trading volume of calls}} \tag{3}$$

$$SI_{S4,t} = \text{Buy-Sell Imbalance } (BSI_t) = \frac{B_t - S_t}{ABS_t} \tag{4}$$

$$SI_{F1,t} = \text{Futures Open Interest Ratio} = \frac{Open_t - \min(Open_t)}{\max(Open_t) - \min(Open_t)} \tag{5}$$

where, B_t (S_t) denotes the buy (sell) value of institutional investors on day t and ABS_t is the average of B_t–S_t, with BSI_t being positive (negative) when the investor group buys (sells) more securities than it sells (buys) on day t. $Open_t$ is the open interest position on day t, and max($Open_t$) and min($Open_t$) refer to the maximum and minimum positions over the sample period. Finally, the put-call ratio measures the sentiment amongst participants in the options market. Participants in a bearish market buy put options to hedge their spot positions or to engage in bearish speculation. The put-call ratio will be higher when investors become more bearish and speculation in puts becomes excessive.

In order to avoid the noise of inflation within the market returns, in this study, we use excess profits and market returns, less the risk-free rate, to analyze both investor sentiment and the futures market. We use a day-of-the-week dummy variable in the following regression to examine the effects of excess profits and to consider the day-of-the-week effect.

The dates presents the descriptive statistics for $\mu_{i,t}$ in the TX, TXF and TXE markets, with all of the markets being leptokurtic and also displaying excess kurtosis. It is necessary to adopt the GARCH model for the distributions, which effectively encompasses the features of asymmetry and fat tails.

We use the aggregate-shock (AS) model, initially introduced by Lin et al. (1994), and subsequently developed by Baur and Jung (2006), with our model accounting for investor sentiment spillover effects between the spot and futures market.

4. EMPIRICAL RESULTS

We begin by estimating the relationship between investor sentiment and excess returns (μ_t). Many of the prior studies have tested whether sentiment can forecast returns or volatility, and it is clear from the responses to the sentiment changes in these studies that sentiment does indeed affect both subsequent returns and volatility.

A positive and statistically significant relationship is found at the 1 per cent level between spot investor sentiment and the returns from both the spot market and the futures market, indicating that stock investor sentiment is capable of raising returns on days following a bullish market. DeLong et al. (1990) argue that the price pressure effect reduces the relative returns expected by noise traders, which suggests that as noise traders become more bullish, they demand more of the risky assets, thereby driving up the price.

The pure mean model is used in this study in an attempt to gain a complete understanding of the relationship between returns and sentiment. The model reveals a positive and statistically significant relationship between futures returns and stock market investor sentiment. The pure volatility model which aims to provide an understanding of the relationship between volatility and sentiment. As expected, stock returns are affected by investor sentiment in the TX, TXF and TXE markets.

Finally, asymmetric volatility is examined in this study based on changes in investor sentiment; this is captured using a dummy variable.A rise or fall in investor sentiment in the stock market clearly does affect futures volatility, particularly in the TX market. As the table shows, individual investors in the stock market can react through either bullish or bearish sentiment; thus, it is of some significant importance to be able to identify and then control for investor sentiment in order to establish a reliable stock timing strategy.

5. SUMMARY AND CONCLUSIONS

Using GARCH models, this study examines the spillover effects of investor psychology on different trading indices. Our results clearly suggest the existence of dynamic relationships between investor sentiment in the spot and futures markets. In particular, we demonstrate that in a bullish market, virtually all stock investor sentiment can result in raising returns; however, no significant relationship is found between futures market investor sentiment and returns. This may be due to the relatively weaker explanatory power of the sentiment index; an issue which may be worthy of future study.

No major differences are discernible in the results obtained from either the pure mean model or the pure volatility model. Furthermore, both the returns in the futures market and investor sentiment in the spot market are found to be positive and significant, which may well suggest that investor sentiment in the spot market has more obvious spillover effects.

Finally, asymmetric volatility attributable to sentiment is found to exist within both the futures and spot markets, whilst sentiment may have an overall effect in the form of either a bullish or bearish reaction. In conclusion, our empirical findings of the spillover effects of sentiment between the spot market and the futures market may be analyzed in terms of their overall implications on market efficiency. Our findings clearly reveal that the links between investor sentiment and market returns are only short-term.

REFERENCES

[1] Antoniou, A., P. Holmes and R. Priestly (1998), 'The Effects of Stock Index Futures Trading on Stock Index Volatility: An Analysis of the Asymmetric Response of Volatility to News', *Journal of Futures Markets*, **18**: 151-66.

[2] Charoenrook and Anchada (2005), 'Does Sentiment Matter?', *Working paper*, Vanderbilt University.

[3] Clarke, R.G. and M. Statman (1998), 'Bullish or Bearish?', *Financial Analysts Journal*, **54**: 63-72.

[4] DeLong, J.B., A. Shleifer, L. Summers and R.J. Waldmann (1990), 'Positive Feedback Investment Strategies and Destabilizing Rational Speculation', *Journal of Finance*, **45**: 375-95.

[5] Fisher, K.L. and M. Statman (2000), 'Investor Sentiment and Stock Returns', *Financial Analysts Journal*, **56**:16-23.

[6] Kock and Koch (1987), 'The Temporal Price Relationship between S&P 500 Futures and the S&P 500 Index', *Journal of Finance*, **42**: 1309-29.

[7] Kumar, A. and C.M.C. Lee (2006), 'Retail Investor Sentiment and Return Co-movements', *Journal of Finance*, **56**(5): 2451-86.

[8] Lee, W.Y., C.X. Jiang and D.C. Indro (2002), 'Stock Market Volatility, Excess Returns and the Role of Investor Sentiment', *Journal of Banking and Finance*, **26**: 2277-99.

[9] Shefrin, H. and M. Statman (2000), 'Behavioral Portfolio Theory', *Journal of Financial and Quantitative Analysis*, **35**(2): 127-51.

A STUDY OF THE RELATIONSHIP BETWEEN KNOWLEDGE-BASED RESOURCES, ORGANIZATIONAL LEARNING AND DYNAMIC CAPABILITIES －THE CASE OF SOLAR ENERGY AND BIOTECHNOLOGY FIRMS IN TAIWAN

Dr. Yuh-Yuan Tsai, Professor, National Dong Hwa University, Taiwan, 03-8633023
Email: yytsai@mail.ndhu.edu.tw
Jia-Rong Jhang, Project Assistant, National Taiwan Normal University, 0912-273616
Email: psybill520@hotmail.com

ABSTRACT

The exploration and exploitation approach in konwledge management has been a focus of interest in the fields of technology management, technology transfer, organizational learning and organizational adaptation. Exploration implies firm behaviors characterized by search, discovery, experimentation, risk taking and innovation, while exploitation implies firm behaviors characterized by refinement, implementation, efficiency, production and selection (March, 1991). When a new technology emerges and reveals uncertain opportunity, a firm may choose to exploit the existing technology for short term survival or to explore the new technology for long term competitive advantage. A strategic choice arises between the balance of exploration of new opportunities and exploitation of existing opportunities, which has been a fundalmental issue for organizational adaptation. The theoretical literature concludes two contrasting arguments to deal with this dilemma: a trade-off or an ambidexterity. These two concepts are used as measurement for organizational learning to investigate the relationship between knowledge-based resources, organizational learning and dynamic capabilities of the emerging industry in Taiwan. The firms of solar energy and biotechnology are taken as the cases for this study.

INTRODUCTION

The whole world is facing the crisis of energy and food shortage. Therefore, several countries began to innovate and develop the solar energy as a substitute for the existing energy resources. Meanwhile, the enterprises in Taiwan have also invested in this industry. Accordingly, the growth rate of output value in biotechnology industry is around 10% in 2008 compared with the previous year. The investments represent the competition status in these two industries. This study attempts to investigate how enterprises respond to the environment change if they use different activities of organization learning and how it relates to the sense of opportunity, the seizure of opportunity and reconfiguration of their existing resources.

Literature points out that knowledge management and organization learning are the important sources of dynamic capabilities. Enterprises could transform their individual knowledge into organizational knowledge by organizational learning. Moreover, they could share, reuse and learn how to adapt to the environment change. The objective of this study is to investigate the relationship among exploration and exploitation of organizational learning, the technical knowledge and market knowledge of knowledge-based resource, and whether the knowledge-based resource and organizational learning would facilitate firm’s dynamic capability

KNOWLEDGE-BASED RESOURCE

In order to articulate the capability-building mechanism, it is first necessary to define the distinction between the terms of ‘resource’ and ‘capability.’ There are two key features that distinguish a capability form other types of resources: First, a capability is firm-specific since it is embedded in the organization and its processes, while an ordinary resource is not. The second feature is that primary purpose of a capability is to enhance the productivity of the other resources that the firm possesses (Makadok, 2001).In order to articulate the capability-building mechanism, it is first necessary to define the distinction between the terms ‘resource’ and ‘capability.’ Capabilities refer to a firm’s capability to deploy resources, usually in combination, using organizational processes, to effect a desired end. They are information-based, tangible or intangible processes that are firm-specific and are developed over time through complex interactions among firm’s resources. They can abstractly be thought of as ‘intermediate goods’ generated by the firm to provide enhanced productivity of its resources, as well as strategic flexibility and protection for its final product or service.

There are two key features that distinguish a capability from other types of resources: first, a capability is firm-specific since it is embedded in the organization and its processes, while and ordinary resource is not. Because of this embeddedness, ownership of a capability can not easily be transferred from one organization to another without also transferring ownership of the organization itself, or some reasonably self-contained subunit of the organization. The second feature is to enhance the productivity of the other resources that the firm possesses. Based on this definition, capabilities cannot easily be bought, they must be built. (Makadok, 2001)

A resource refers to an asset or input to production (tangible or intangible) that an organization owns, controls, or has access to on a semi-permanent basis. An organizational capability refers to the ability to a coordinated set of tasks, utilizing organizational resources, for the purpose of achieving a particular end result (Helfat & Peteraf, 2003). Capabilities, in contrast, refer to a firm’s capability to deploy resources, usually in combination, using organizational processes, to affect a desired end. They are information-based, intangible or intangible processes that are firm-specific and are developed over time through complex interactions among the firm’s resources. Capabilities are often developed in functional area (e.g., brand management in marketing) or by combining physical, human, and technological resources at the corporate level (Amit & Schoemaker, 1993). In addition, resources and capabilities are used interchangeably and refer to the tangible and intangible assets firms use to develop and implement their strategies (Ray, Barney & Muhanna, 2004).

Resource-based theories of strategy (RBV) argue that firms with valuable, rare, inimitable and nonsubstitutability resources have the potential of achieving superior performance (Barney, 1991, 1995). Resources are inputs into a firm’s production process and can be separated into those that are knowledge-based and those that are property-based. Property-based resources typically refer to tangible input resources, whereas knowledge-based resources are the ways in which firms combine and transform these tangible input resources. Knowledge-based resources may be particularly important for providing sustainable competitive advantage, because they are inherently difficult to imitate, thus facilitating sustainable differentiation, play an essential role in the firm’s ability to entrepreneurial and improve performance. Knowledge about markets and technology represent two strands of procedural knowledge that potentially have strong performance implications, because they increase the ability to discover and exploit opportunities. Market knowledge can increase a firm’s ability to discover and exploit opportunities because: (1) awareness of customer problems may have great generality and constitute real market opportunities; (2) it is easier to determine the market value of new scientific discoveries, technological change etc.; (3) the locus of innovation often lies with users of new technologies who cannot easily articulate their needs for not-yet-developed solutions to problems, and therefore the organization must share some of the same tacit knowledge as its users. Technological knowledge can also enhance the discovery and exploitation of opportunities and provides a firm with the ability to rapidly exploit opportunities, or to be able to respond quickly when competitors make advancements (Wiklund and Shepherd, 2003)

ORGANIZATIONAL LEARNING

March (1991) refers two types of learning to exploration and exploitation. Exploration includes things captured by terms such as search, variation, risk taking, experimentation, play, flexibility, discovery, innovation. Exploitation includes things as refinement, choice, production, efficiency, selection, implementation, execution. Organizational exploration is search for new knowledge, use of unfamiliar technologies, and creation of products with unknown demand. Because these activities do not reliably and quickly produce revenue, exploration has uncertain and distant benefits. Exploitation is use and refinement of existing knowledge, technologies, and products, and has more certain and proximate benefits. Exploration and exploitation both draw resources, and thus resource constraints require organizations to make tradeoffs between them (Levinthal & March, 1993).

Most researchers agreed with the perspective of March (1991) and believe that exploration and exploitation are the two ends of a continuum and compete with each other for the limited resources within an organization (Ozsomer & Gencturk,2003; Holmqvist,2004; Auh & Menguc,2005; Lavie & Rosenkopf, 2006). However, the other researchers argued that some resources are unlimited, such as knowledge, information, and external resources (Benner & Tushman,2003; He & Wong,2004). They believe that there is no contradiction between explorative and exploitative activities: The phenomena are labeled "orthogonality" versus "continuity" by Gupta, Smith, and Shalley (2006). The former believe that firms make decision for the ratio of resource allocating to explorative and exploitative activities according to the strategic position, they chose, within the continuum, while the latter believe that explorative and exploitative activities are complement with each other. Researchers, who believe that exploration and exploitation are two different and orthogonal aspects of organizational behavior, note that firms may acquire external resources by engaging in alliances, mergers, and acquisitions. Cesaroni, Minin, and Piccaluga (2005) observe firms' behaviors under the situations where technological changes are incremental or fairly rapid. They found that when technological change is fairly rapid and the dominant technology is not formed, focusing on R&D activities or looking for technological cooperation my help firms to survive. On the other had, when technological change is incremental and the firm is short of complementary asset, it can use its alliances' assets to exploit the outcome of its innovation.

While researchers have normally assumed that firms should seek to balance exploration and exploitation because both short-term productivity and long-term innovation are essential for organizational success and survival (March, 1991; 87), Lavie and Rosenkopf (2006) found that firms have different tendencies within the exploration-exploitation continuum in different domains which are complement with one another.

Birkinshaw and Gibson (2004) proposed concepts of "adaptability" and "alignment" which are corresponsive with exploration and exploitation respectively. Adaptability represents the capability to move to the new opportunity for the adoption of changing markets and creating long-term value.

Alignment represents the capability to coordinate different value activities and concentrate on short-term value creation. If a firm want to acquire long-term benefits, it have to foster both capabilities---The firm is labeled "ambidextrous organization" (Birkinshaw & Gibson, 2004). Firms which focus on alignment excessively will enjoy profits promptly, but will reduce the capabilities to adapt to the change of the environment in the future. Similarly, too much attention to the adaptabilities side of the equation means building tomorrow's business at the expense of today's and will make shareholders unsatisfied. It is a challenge for firms to balance both capabilities within limited resources.

DYNAMIC CAPABILITY

Teece et al.(1997) defines dynamic capability as the firm's processes that use resources specifically the processes to integrate, reconfigure, gain and release resources to match and even create market change. Dynamic capability thus is the organizational and strategic routines by which firms achieve new resource configurations as markets emerge, collide, split, evolve, and die. A firm's dynamic capability is determined in three categories: processes, positions, and paths. Processes refer to the way things are done in the firm, or what might be referred to as its routines, or patterns of current practice and learning. Positions refer to its current endowment of technology and intellectual property, as well as its customer base and upstream relations with suppliers. Paths refer to the strategic alternatives available to the firm, and the attractiveness of the opportunities which lie ahead. The firm's processes and positions collectively include its capability or competence.

Luo (2000) defined dynamic capability as a multinational enterprise's (MNE) ability to create, deploy, and upgrade organizationally embedded and return-generating resources in pursuit of sustained competitive advantage in the global market place. In this view, dynamic capability requires a strong base of established capability or resources as well as the ability to efficiently deploy these resources and to continuously create bundles of new resources and knowledge. Teece and Pisano (1994) summarize the components of dynamic capability framework as (1) organizational and managerial processes: coordination/integration, learning and reconfiguration and transformation; (2) positions: technological assets, complementary assets, financial assets, reputation assets, structural assets, institutional assets, market (structure) assets and organizational boundaries; (3) paths: path dependencies and technological opportunities. Based on above perspectives, this study proposes 3 categories such as core competence, organizational capability and capability upgrading for a firm's dynamic capability.

For analysis purpose, dynamic capabilities can be disaggregated into the capability (1) to sense and shape opportunities and threats, (2) to seize opportunities, and (3) to maintain competitiveness through enhancing, combing, protecting, and reconfiguring the business enterprise's intangible and tangible assets. Dynamic capabilities include difficult-to-replicate enterprise capabilities required to adapt to changing customer and technological opportunities (Teece, 2007).

RESEARCH FRAMEWORK AND HYPOTHESES

The proposed research framework, shown in Figure 1, enumerates knowledge-based resource affecting dynamic capability with organizational learning as intervening construct. Within this framework, dynamic capability of sensing, seizing and transformation is affected by exploration and exploitation.

Meanwhile, organization learning is affected by knowledge-based resource of technology knowledge and market knowledge.

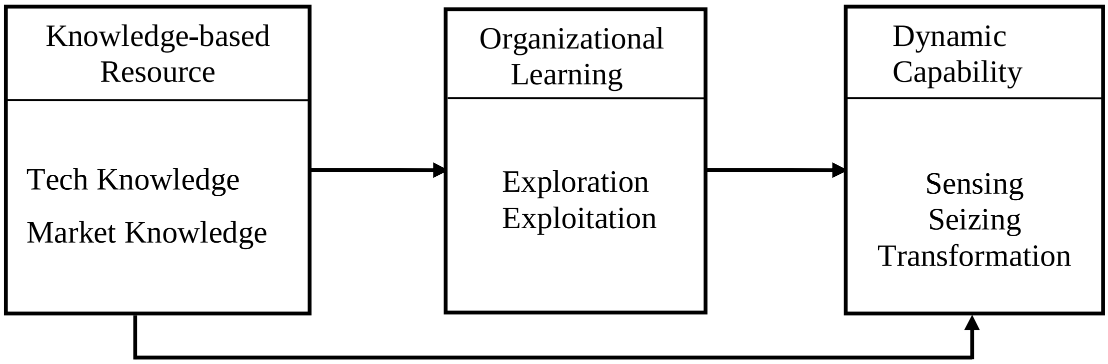


Figure 1: Research Framework

The hypotheses are proposed as follows by literature review and in-depth interviews:
Hypothesis 1: The level of knowledge is positively related to organizational learning.
Hypothesis 2: The level of organizational learning is positively related to its dynamic capability.

MULTI-REGRESSION ANALYSIS AND HYPOTHESIS TEST

The results of multi-regression analysis for solar energy industry and biotechnology are illustrated in Table 1 and Table 2 as follows:

Table 1 Multi-regression analysis for solar energy industry

Dependent variables	independent variables	β	P	VIF
exploration	technology knowledge	-0.016	.894	1.103
	market knowledge	4.25	.001**	1.103
exploitation	technology knowledge	0.484	.000***	1.103
	market knowledge	0.277	.01*	1.103
sensing	technology knowledge	0.114	.269	1.103
	market knowledge	0.606	.000***	1.103
seizing	technology knowledge	0.247	.045*	1.103
	market knowledge	0.203	.110	1.103
transformation	technology knowledge	0.206	.110	1.103
	market knowledge	0.208	.105	1.103
sensing	exploration	0.472	.000***	1.510
	exploitation	0.290	.014*	1.510
seizing	exploration	0.177	.221	1.510
	exploitation	0.285	.044*	1.510
transformation	exploration	0.232	.113	1.510

	exploitation	0.219	.133	1.510
sensing	R^2 = 0.599，adj R^2 = 0.572			
seizing	R^2 = 0.202，adj R^2 = 0.149			
transformation	R^2 = 0.188，adj R^2 = 0.133			

*indicates P<.05 , **indicates P<.01 ,***indicates P<.001

Table 2 Multi-regression analysis for biotechnology industry

Dependent variables	independent variables	β	P	VIF
exploration	technology knowledge	.001	.995	1.047
	market knowledge	.279	.019*	1.047
exploitation	technology knowledge	.380	.001**	1.047
	market knowledge	.241	.025*	1.047
sensing	technology knowledge	.217	.065	1.047
	market knowledge	.156	.182	1.047
seizing	technology knowledge	.525	.000***	1.047
	market knowledge	.134	.183	1.047
transformation	technology knowledge	.467	.000***	1.047
	market knowledge	.191	.066	1.047
sensing	exploration	.084	.593	2.070
	exploitation	.340	.033*	2.070
seizing	exploration	-.278	.064	2.070
	exploitation	.644	.000***	2.070
transformation	exploration	-.146	.345	2.070
	exploitation	.526	.001**	2.070
sensing	R^2 = 0.181，adj R^2 = 0.134			
seizing	R^2 = 0.376，adj R^2 = 0.339			
transformation	R^2 = 0.326，adj R^2 = 0.287			

*indicates P<.05 , **indicates P<.01 ,***indicates P<.001

CONCLUSION

According to the questionnaires survey and its statistical analysis in solar and biotechnology industries, we found that: (1) The organizational learning has positive effect on the dynamic capability, (2) The knowledge-based resource has positive effect on the organizational learning, (3) The knowledge-based resource has positive effect on the dynamic capability, and (4) there are significant differences between knowledge-based resource, organizational learning and dynamic capability in these two industries. Specifically speaking, market knowledge will enhance explorative learning while technology knowledge will enhance exploitative learning and which will accordingly improve their dynamic capability.

UNDERSTANDING DYNAMIC CAPABILITIES BY TRANSFORMATION OF ORGANIZATIONAL SLACK

Ching-Chou Chen, National Dong Hwa University, (886) 3863-3017, ching@mail.ndhu.edu.tw
Hsin-Hua Hsiung, National Dong Hwa University, (886) 3863-3027, hsiung@mail.ndhu.edu.tw

ABSTRACT

This research proposes a new view on dynamic capabilities, which differs from prior studies which emphasized the types of capabilities. In this study, dynamic capabilities are viewed as a feedback process of how organizations can transform slack into particular innovations, which in turn can influence organizational performance. Organizational innovations in an intermediate role can mediate the relationship between organizational slack and performance. The transformation within the process may be moderated by core competency (internal factors) and environmental turbulence (external factors), sequentially. Simultaneously, it can also solve the problem of an unclear relationship between slack and performance through the understanding of dynamic capabilities in such a way.

INTRODUCTION

The more turbulent the environment confronted by organizations the more dynamically the capabilities will be viewed as a kind of critical resource base by strategic management. Though the concept of dynamic capabilities, it has been posited that an organization can display the performance impact of existing resources through specifically the processes to integrate, reconfigure, gain and release resources (Teece, Pisano, & Shuen, 1997). Many studies have different views of what dynamic capabilities are, such as differences in the scope of related diversification (Doving & Gooderham, 2008), resource management capability (Moliterno & Wiersema, 2007), victory over potential rigidities of organizational capability building (Schreyogg & Kliesch-Eberl, 2007), R&D and marketing resource deployments (Kor & Mahoney, 2005), and a set of routines guiding the evolution of a firm's resource configuration (Zott, 2003). We can observe the function of dynamic capabilities from the various capabilities themselves or from their transformation process. Previous research has contributed the functional content of dynamic capabilities in comprehending organizational response on environmental turbulence, especially on diverse capabilities. This study will take the view of transformation process of organizational resources for dynamic capabilities.

There are particular resources to be changed or transformed within an organization. Such kinds of resources may not be strategic resources, which are used in current operations and which are valuable to the competition before transformation. Organizational slack may be an important resource rather than a strategic resource which fits criteria such as value, rarity, imperfection imitability and being without strategically equivalent substitutes (Barney, 1991). To decompose the dynamic capabilities from the transformation of organizational slack provides us with a better theoretical perspective. Prior research in dynamic capabilities has been based upon an assumption that firms' strong capabilities are beneficial in that they quickly respond to environmental turbulence (Eisenhardt & Martin, 2000; Song, Droge, Hanvanich, & Calantone, 2005). However, in this transformation process, what happens in organizations and what changes occur within organizations? Dynamic capabilities may be a function that it changes things which are not strategic resources into strategic resources, and then affects organizational performance under particular environmental conditions. If we can learn more about these processes, we can understand more about dynamic capabilities.

As with other strategic capabilities such as learning capabilities, flexibility and networking, dynamic capabilities also have the nature of being hard to identify. Past research has put too much emphasis on the concept of capabilities themselves in that managers find it hard to work with them in reality. These capabilities have the nature of being hard to discern. Unlike explicit resources, the characteristics of dynamic capabilities are not easily discernible from a quick physical inspection of managers in the way that organizational slack resource characteristics often are. Furthermore, either raising or promoting the capabilities requires spending more time than reallocation in a particular period for organizational slack. From the organizational slack transformation, we can apparently map out the energy of dynamic capabilities. In addition, all of the above studies presume the same functions which change an existent position or transform current resources and reveal the same factors which are internal capability and external environment. The research gap this study filled is that the transformation process of visible resources can increase understanding of dynamic capabilities more clearly.

Responding to variable competitive conditions, an organization would have a chance to improve its performance if it did something different to what the others did. Organizational slack provides such an opportunity for innovation to change (Voss, Sirdeshmukh, & Voss, 2008). Organizational innovations play the pivotal role between organizational slack and performance. Innovations have changed and improved existing products, markets, finance and operational rules. Slack facilitates firstly innovations which then improve sequential organizational performance. Whether an organization is able to make slack into innovation depends on its internal capabilities. Internal capabilities include organizational learning, re-configurability and coordination capability, which belong to a kind of dynamic capability within organizations. It depends, then, on external environmental turbulence whether innovations can increase organizational performance.

From the slack transformation process, we decompose dynamic capabilities into a two-stage process, slack-innovation and innovation-performance. In the first stage, interaction between slack and core competency affects innovation. In the second stage, interaction between innovation and environmental turbulence affects performance. Owing to these critical components, we construct their relationship and then let us clearly make sense of the function of dynamic capabilities.

RECYCLING ORGANIZATIONAL SLACK

One of the explicit operations for dynamic capabilities is the organizational process or routine by which the firm configures and reconfigures its portfolio of strategically and non-strategically important resources (Eisenhardt & Martin, 2000; Helfat, 1997; Moliterno & Wiersema, 2007; Teece et al., 1997; Winter, 2003). For these critical resources, slack is one particular resource that can be reallocated, which is consistent with the direction of firm's strategies (Barney, 1991; Dierickx & Cool, 1989; Mishina, Pollock, & Porac, 2004). Slack has been defined as the pool of resources within organizations that is in excess of the minimum necessary to produce a given level of organizational output (Nohria & Gulati, 1996). Organizational slack may be produced from managerial cognition (Danneels, 2008) and innovation may be one kind of outcome of strategic action response in the environmental context. If we could connect the created artificial boundaries, we would not limit our ability to develop holistic explanations for strategic action (Nadkarni & Barr, 2008).

Our research builds on prior research, investigating various aspects of dynamic capabilities to articulate a more comprehensive understanding of the relationship between slack and performance. If dynamic capabilities could be broken down into several stages, the process, especially, would make the function more clarified. Teece (2007) suggested that dynamic capabilities can be disaggregated into the capability (1) to sense and shape opportunities and threats, (2) to seize opportunities, and (3) to maintain competitiveness through enhancing, combining, protecting, and, when necessary, reconfiguring the business enterprise's intangible and tangible assets. Similarly, Schreyogg and Kliesch-Eberl (2007) note that capability monitoring exercises at an observational level for capability practice on an operational level interact with both the internal and external environment. Basically, if organizations can sense and monitor the internal and external change, they will have certain expectations and preparations for the next period of environmental turbulence.

Organizational slack can be produced by the results of pre-period strategic decisions, real conditions of which may fit managers' expectations or not (Mishina et al., 2004). In this sense, organizational slack has neutral characteristics and whether it is good for organizations or not depends on core competency. If organizations possess sufficient and necessary capabilities, they can transform slack into innovations and fit their environmental needs. However, it is redundant for organizations that do not have enough capabilities to transform slack into innovations or judge their environmental needs. The more environmental turbulence there is, the harder it is for the managers to exactly predict how to judge their needs.

The point to observe is that organizational slack can stimulate the production of innovations, which in turn can affect organizational performance. The performance feeds back into intended slack. Even though there is unintended slack when actual variables are out of expectation, the firms under pressure still need to transform all of that into organizational innovations if they own superior core competency.

The source of organizational slack

Resource allocation and reallocation make organizations flexible in response to environmental impacts (Levinthal & Posen, 2007; Miller, Lant, Milliken, & Korn, 1996). Decision makers in organizations have expectations for the next period of environmental turbulence before reassigning critical resources. In general, the unplanned or planned results of the decisions would become one of three conditions: over, under, and fit expectations. But according to Voss, Sirdeshmukh, and Voss (2008), there are four types of slack resources: financial, operational, human resource and customer relational slack. We connect their relationship. Firstly, preparation over expectation, where there is the formation of operational slack and human resource slack. Secondly, preparation matches expectation, where there is no accumulation of operational and human resource slack but one of financial slack. Thirdly, preparation under expectation, where there is no accumulation of operational slack but a customer relational slack. Finally, there is an additional case of planned results, where organizations purposely retain some slack with which prepare to take particular strategic activities that sustain the competitive advantage.

The effects of organizational slack

If organizational slack is idle in a particular firm, this shows that the firm does not rapidly have the capability in responding to environmental shifts (Cheng & Kesner, 1997) or strategic intent (Julian & Ofori-Dankwa, 2008) and will subsequently fail in its performance. Strategic mangers must take action if slack is produced actively or passively. Organizational slack can produce either planned or unplanned results. So, slack is a dynamic quantity that represents the difference between the resources currently possessed by the firm and the resource demands of their current business (Mishina et al., 2004). Organizational slack may become a kind of pressure to change whether within or beyond organizations. The more dynamic capabilities organizations have, the more organizations can transform slack into innovations. Organizational innovations are defined as new or significantly amended forms of organization, business structures or practices, aimed at step changes in internal efficiency of effectiveness or in approaching markets and customers (Battisti & Stoneman, 2009).

In sum, the components of this theory of dynamic capabilities are organizational slack, innovations, performance, core competency, and environmental turbulence. Each of these leads to propositions on the feedback process of dynamic capabilities, which are developed next.

INTERMEDIATE ROLE OF INNOVATIONS

In the field of strategy management, the relationship between organizational slack and performance is always an unsolved problem (Daniel, Lohrke, Fornaciari, & Turner, 2004; George, 2005; Tan & Peng, 2003). There are some undiscovered processes and means within the relationship. From the perspective of dynamic capabilities, organizations have the ability to transform slack into valuable innovations in response to environmental turbulence under the expectations of the environment. If the managers' expectations match the reality of the environment's turbulence, innovations have values for organizations and will increase organizational performance. Organizational innovation is defined as adoption of an internally generated or purchased device, system, policy, program, process, product, or service that is new to the adopting organization (Damanpour, 1991). In this study, we assume the definition of outcomes for innovations.

Slack and Innovations

Though organizational slack may come from intended or unintended results, the firm always needs to find a new way that makes beneficial changes or value-added outcomes. While active innovation may search for better opportunities for profits, passive innovation may bring pressure to improve their management. Change for organizations is often to make innovations (Cheng & Kesner, 1997). Innovations refer to risk taking and doing something new within organizations, such as entering a new market, promoting the quality of employees, improving production processes, and satisfying particular segmental customers.

Organizations would be taking a risk to innovate either by a problem-driven search or a slack-driven search according to management literature (Baum, Rowley, Shipilov, & Chuang, 2005; Cyert & March, 1963; Greve, 2003). Even though initial ideas are driven by problem-solving, organizations may need to create slack in order to innovate so that firms may reallocate their resources. That is, in the case of a problem-driven search, firms will firstly create slack for innovations. If there is no slack, there will be no room to innovate. Indeed, Rosner (1968) concluded that the existence of slack means that the organization can afford (1) to purchase costly innovations, (2) to absorb failures, (3) to bear the costs of instituting the innovation and (4) to explore new ideas in advance of an actual need. We expect that slack will increase the potential for innovations:

Proposition 1: Organizational slack will increase innovations in organizations.

Innovations and performance

Organizational managers could have some expectations about their innovations before making them. In general, because the environmental turbulence is more rapid than before, the firm must be more flexible in response (Lawless & Anderson, 1996). That organizational innovation comes from pre-period preparations is the flexibility to change the firms themselves in response to the environmental turbulence (Voss et al., 2008). Organizational innovations may represent adaptability to change. The more innovations firms make, the more organizational performance firms gain. This gives rise to a second proposition:

Proposition 2: Organizational innovations will increase performance.

TAKE ACTION FOR SLACK

Organizational slack has to pass an internal process and an external process and then affects performance. Someone executing the process of dynamic capabilities needs to consider existing internal and external conditions. Relatively unchangeable components in the process are core competency for internal factors and environmental turbulence for external factors. Organizational slack could be a function of performance, which needs to pass two phases. In the first phase, slack could be successfully transformed into innovation through internal conditions (i.e. core competency). In the second phase, organizational slack could promote performance through external conditions (i.e. environmental turbulence). The organizational environment will affect both the use of slack at this time and the final utility of slack the next time (Voss et al., 2008).

Internal Conditions

Organizations often fail to realize the potential benefits of innovations. One of the important reasons is lag. Organizational lag refers to the discrepancy in the rate at which new technical and administrative ideas are implemented (Damanpour & Evan, 1984). A balanced rate of adoption between different innovations is more effective in helping organizations to maintain or improve their level of performance than either particular innovation alone; for example, administrative innovations could frequently lag behind technical innovations in an organization and that leads to low performance (Daft, 1978; Damanpour & Evan, 1984). Organizational lag implies that firms cannot follow up on the environmental dynamics and not possess enough capability to reallocate organizational slack. There may also be a kind of signal for organizational inertia so that firms cannot adjust their strategic directions. In other cases, there may be a preparation for new strategic activities that firms intend to produce. Whether the change may be an opportunity or a threat for the organization depends on organizational capabilities taking adequate strategic directions for innovations. If the firms have superior capabilities, especially core competency, the slack activity that will be rapidly transformed into innovations would be more beneficial. This leads to another proposition:

Proposition 3: The effects of organizational slack on innovations will be moderated by core competency; organizational slack is more likely to increase innovations when core competency is high than when it is low.

External Conditions

Another reason organizations often fail to realize the benefit of innovations is that there is confusion between goals and reality; in other words whether the direction of organizational goals is really similar to the real environment. Though the firms' managers may realize that the environment is turbulent, they will not be able to predict when a turbulent situation will happen, nor exactly know its magnitude, direction, and category. There are opportunities for organizational innovations if the environment has changed and has yet to match others expectations (Pierce, 2009).

Third, responsive fitness, which refers to firms' responsiveness to the environmental turbulence, is neither over or underestimated. There are threats for organizational innovations if the environment has too much turbulence as compared to what they anticipated, so that the innovations will be less useful

(Nadkarni & Barr, 2008). The innovations fitting anticipated change and preparations will increase organizational performance, but innovations that do not fit change will be redundant for organizations. This leads to a fourth proposition:

Proposition 4: The effects of organizational innovations on performance will be moderated by environmental turbulence; organizational innovations are more likely to increase performance when environmental turbulence is low than when it is high.

Based upon the above evidence from the literature review, we propose a feedback process model of the three steps of dynamic capabilities: organizational slack, organizational innovations, and performance, as diagrammed in Figure 1.

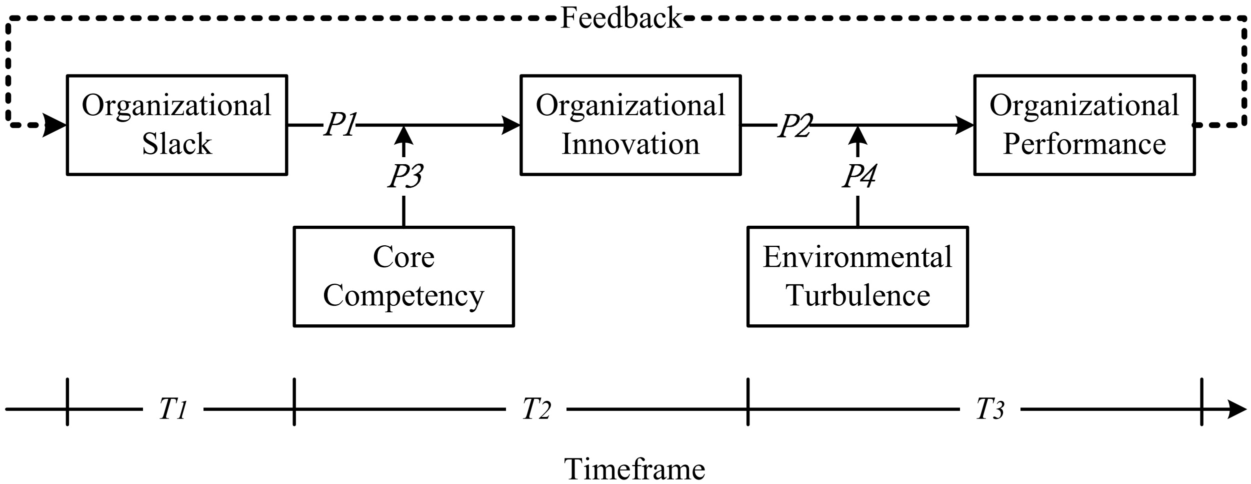


Figure 1. The feedback process model of dynamic capabilities from organizational slack transformation

CONCLUSION

Regarding dynamic capabilities, researchers need to consider internal and external factors in organizations. The mechanism of dynamic capabilities not only includes capability itself, but also the process of resource transformation, such as slack resources and innovations.

There is now a better way to understand dynamic capabilities from resource transformation. We view dynamic capabilities as a feedback process that functions to transform slack into innovations and which then affects performance. Based on the strategic management literature, whether inside out or outside in strategic views (Hoskisson, Hitt, Wan, & Yiu, 1999), there is a loop cycle between slack, innovations and performance. This study begins with organizational slack and explores the process to performance.

There may not be a so called "paradox of organizational capabilities (Schreyogg & Kliesch-Eberl, 2007)," but rather a problem of whether the reality fits the managers' expectations or not. Organizations could reallocate slack at this time and repeatedly and continually adjust their innovations at a later time in response to the needs of the environment.

REFERENCES

References available upon request from the first author.

The Practice of Internal Marketing is Influence the Market Orientation and Innovation

Yafang Tsai
Associate Professor, Department of Health Policy and Management,
Chung-Shan Medical University, Taiwan
+886-4-24730022 ext.17176, avon611@yahoo.com.tw

Shih-Wang Wu
Assistant Professor, Dept. of Hospital and Health Care Administration,
Chia Nan University of Pharmacy & Science, Taiwan
PhD Candidate, Graduate Institute of Business Administration,
National Chung Cheng University, Taiwan
+886-6-2664911 ext.5225, scottwu101@mail.chna.edu.tw

Hsien-Jui Chung
Professor, Dept. of Business Administration,
National Chung Cheng University, Taiwan
+886-5-2720411 ext.34313 , hjchung@ccu.edu.tw

ABSTRACT

The purpose of this study is to explore that how hospitals' administrators apply the practice of internal marketing to influences hospitals' market orientation and innovation. The research subjects are 491 contracted hospitals with the evaluated level as district hospital and above to investigate presidents and high-level managers of hospitals in Taiwan. There were 106 valid questionnaires retrieved and the response rate was 21%.According to the result, this study found that: Internal marketing influence the market orientation significantly; Internal marketing influence the organizational innovation significantly.

Keywords: Hospital management, Healthcare service, Market orientation, Innovation, Internal marketing

INTRODUCTION

The Taiwanese National Health Insurance Program has been successfully implemented over the years however just like the hospitals in countries like Europe or the United States, they faced increasing market competitions and managerial challenges upon implementing insurance pay out program. In order to sustain its competitive advantages, standing out within the Medical market, over the past years hospital administrators in Taiwan have devoted their effort in improving internal control systems, hopping through better medical service procedure and well maintained doctor patient relationship, will build a long term partnership with patients in securing stable patient source.

The objective for a good medical care is to create value attended to patients' needs (Porter & Teisberg, 2006) .The basis of this value creation is to satisfy patients' needs (Quinn et al., 2009). This is to provide suitable medical service centered on patients' needs. The internal marketing department will provide a specific vision to all employees, through establishing internal education and training which will strengthen employees' service ability and efficiency where further improving overall service quality ultimately achieving customer satisfaction. Therefore, for this reason internal marketing is crucial especially to medical institutions (Tsai & Tang, 2008).

Besides using practical internal marketing in improving employee's service ability, hospital administrators needed to assign a clear service objective. With the current advancement in medical technology and technique, hospitals will strive for continuous innovation, improving the technical knowhow of medical personals or at the same time introducing state of the art equipments so to provide stable medical service quality (Lega, 2009) in order to maintain long term patient relationship. This will not only generate stable income for the hospitals but also build up a strong competitive advantage against other competing hospitals.

BACKGROUND

Internal Marketing

The implementation of internal marketing within hospitals not only gives employees a clear working target (Boshoff & Tait, 1996) but can also used in resolving employee's working needs and desire. Through internal practical marketing, organisation can increase employee's work satisfaction furthermore team members will also be taking a more positive role in completing tasks assigned by institution (Bernstein, 2005). Internal administrators applied practical human resource practice; motivate, mobilize, co-operation and manage to understand the working needs of internal customer(Employees), then by satisfying those working needs employees will be happy to contribute to the organization (Joseph, 1996)

The frontline employee plays a vital role in the process of delivering service (Congram & Friedman, 1991). Organisation's long term vision is communicated to the frontline staff allowing the frontline staff to identify themselves with the vision at the same time influencing employees' service behavior so as to improve the service quality provided to the external customers. Tsai & Tang (2008) applied Factor Analysis that classified hospital's internal marketing practice into 3 different programs, Service training programmes, Performance incentives and Vision about excellence service. They discovered that among these 3 programs, Service training programme and Vision about excellence service had influences on employee's attitude towards service quality.

Lara & Mesa (2006) believed that Market Orientation is the most crucial marketing strategy within the medical industry in the 21st century. The concept of Market Orientation is suggested by 2 great schools of thoughts Kohli, & Bernard(1990) and Slater & Narver(1994). They viewed that the traditional marketing concept emphasis heavily on the execution of marketing activities and planning which were too constricted. They felt that organization should be more active in gathering external market intelligence, disseminating those intelligence to members of relative departments hence induce immediate response to this information (Kohli, & Bernard, 1990).

Organisation performance can be improved through switching to a market orientated focus organization culture (Slater & Narver, 1994). This market orientated organization culture is formed by 3 behavioral combinations: customer orientation, competitor orientation and inter-functional coordination. Besides fulfilling the demands and needs of external customers, organization must also keep abreast of any changes in competitors' marketing strategy, passing through market intelligence to members of relative departments then further by cross functional communication and discussion, research and develop a product or services which conform to customers' needs allowing an immediate response to customers' needs by the organization.

The concept of market orientated marketing provided organization with an operational philosophy that the focus of organisation's key coordination activities should be placed on how to satisfy customers' needs (Tay & Tay, 2007), with the aim of creating customer value. For this reason, organizations who possessed market orientated culture will communicate frequently, sharing intelligence relating to customers' needs at the same time respond to those needs (Narver & Slater, 1990). Therefore customer value is created through such efficient and result driven method. In the long term perspective, it would generate returns to the organization (Slater, & Narver, 1994b; Harris, 2001).

Organizational innovation

The effectiveness of bringing innovative concepts into practice by health care personals is an important topic within heath care industry (Mylopoulos & Scardamalia, 2008). Organisation that possessed an innovative culture generally has an open attitude in accepting new concepts (Hurley, & Hult, 1998). With the advancement in technology, health care organizations are forced to adopt innovative concepts (Shekelle et al., 2005). When facing the competitive health care industry, with an aim in continuous innovation, hospitals should build a working environment which encourages creative behaviors within employees.

Innovative environment is centered around effective management in organization's basic infrastructure to support creativity; innovative behavior is centered around employees, the execution innovative activities in organization is lead by the unique behaviors of employees (Dobni, 2006). An organization which possessed an innovative culture should comprise these characteristics including market orientation, innovation-friendly ruling coalition and promoting organization structure (McCosh et al.1998). Within these characteristics promoting organization structure relates to innovative environment. Integration of market orientation and innovation can be achieved through employee behavior, technical knowhow or ability.

Damanpour, & Evan(1984)were in the opinion that innovative management occurred within organisation's community system that include factors such as human resource, reward system, procedure control system and organization structure. Technology innovation refers to products, service and manufacturing technology which interconnects with product or procedures. Organisation that possessed innovative culture should also possess the ability in both innovative management and technical innovation (Kim, & Srivastava ; Gopalakrishnan, & Damanpour, 1997).

The effect of internal marketing to market orientation

Hospital managers practice internal marketing in communicating organisation's vision and mission to the employees then through education and training, service skills are then developed ultimately building a customer focus culture within the organization. Such practice will enhance employee's service attitude towards external customers, ensuring a more efficient delivery of service value to external customers (Brad, 2004) achieving the objective of market orientation.

Market orientated organization will utilize its managing philosophy in providing behavior, attitude and operation guidance to its members (Loubser, 2000) so to increase internal operation efficiency in achieving better results. The application of internal marketing concept is to inspire employees in skills development and by applying those skills within organization (Grönroos, 1985), achieving external customer satisfaction through improved service efficiency. In order to fulfill the concept of customer orientation, organizations who value this concept will communicate this idea with internal customer (employee) (Hopkinson, 2003) conveying organisation's vision and mission to their employees. Organisation's commitment to service quality is then passed through by providing value added service to external customers by its employees. Therefore, the practice of internal marketing within organization will affect the forming of market orientation culture.

Hypothesis 1 Practice of internal marketing within organization will affect the forming of market orientation culture.

The effect of internal marketing to organization innovation

Organisation projects or activities often require co-ordination (Kouvelis & Lariviere, 2000). When market intelligence is gathered through interactions between frontline staffs and customers, this intelligence must be shared and communicated within inter-functional departments (Song, Michael, Mitzi, & Schmidt, 1997) so to transform this information into a practical new product development concept resulting a quicker response to customer demands. The advantage of inter-functional cooperation allows better utilization of knowledge, through brain storming session team members are challenged in formulating new product or service concepts (Boyd, 2007). This is particularly beneficial in new product or service exploitation or creation. Internal marketing stressed the importance of building a good internal communication channel by administrators further encouraging cooperation between inter-functional departments allow information and knowledge sharing within members. This will enhance creativity in organization members (Tsai, 2002).

Value added service is often provided by the frontline employee through interaction with the customers (Edvardsson et al., 1997). Market intelligence can be gathered in the course of first point interaction between frontline employee and customers however this information must be transmitted to members of various departments through internal communication allowing better understanding of customer needs. New product, service or new manufacturing process is then created through further inter-functional communication and coordination at the same time current operation strategy may also be amended accordingly (Cheng & Van De Ven 1996). This highlights the benefits of practicing internal marketing in organization innovation.

Hypothesis 2 The application of internal marketing is beneficial to organization innovation.

METHOD

This research adopt a cross sectional study approach, using questionnaires in gathering research data based on certified hospitals who joined The Taiwan National Health Insurance Program between the Year of 2001 to 2005. Researchers posted out in total 491 copies of questionnaires on October 2008. Based on both postal and phone interview in total 106 copies of questionnaires were received, overall collection rate is 21%.

Internal marketing- Measurement of internal marketing is based on factor analysis by Tsai & Tang(2008) after analyzing internal marketing practice can be classified into 3 different programmes: service training programmes (Cronbach's alpha = .92), performance incentives (Cronbach's alpha = .82), and vision about excellence service (Cronbach's alpha = .86), in total 14 questions. Cronbach's alpha is 0.957. Market orientation- This research is based on by Slater & Narver(1994), in total of 14 questions. Based on Likert's 5 point measuring scale, responses are classified as "totally not agreed", "not agreed", "average", "agreed" and "totally agreed". Questionnaire Cronbach's alpha is 0.913.Organisation innovation- Organisation which possessed innovative culture, should also include management and technology innovation (Kim & Srivastava; Gopalakrishnan & Damanpour, 1997). This research is measured according to management and technology innovation, in total of 19 questions. Using Likert's 5 point measuring scale, responses are classified as "totally not agreed", "not agreed", "average", "agreed" and "totally agreed". Questionnaire Cronbach's alpha is 0.958.Research data is analyzed by using SPSS for Window 17.0 Statistics analysis programme.

RESULTS

This research is based on linear regression to exam the hypotheses 1 to 5. According to the research analysis result, it was discovered that (1)Internal marketing has substantial effect on market orientation (R^2=0.347, β=0.589, P=0.000). (2) Internal marketing has substantial effect on organization innovation (R^2=0.664, β=0.815, P=0.000). (3) Service training programme in internal marketing had obvious effect on organization innovation (R^2=0.631, β=0.794, P=0.000).

CONCLUSION AND IMPLICATION

Health care service is a form of behavior which requires interaction between employees and patients then through employee organization’s commitment in its service value is then achieved. Besides emphasizing on satisfying external customer needs, it is also important to value employee’s working needs, by satisfying those needs, gaining employee’s recognition and sense of belonging creating a happier workforce which will commit to the organization. This research found, hospitals that practices internal marketing will affect its attitude towards market orientation, in other words hospitals that implement internal marketing will affect employee’s attitude towards customers. It is recommended to hospital administrators to utilize internal marketing activities, attract and retain outstanding medical professions, increase service capability which will provide positive effect on becoming a customer orientated organization.

Health industry in Taiwan is currently facing a structural transition period for example; enormous operational pressure caused by changes in health care and health insurance pay out policies. In order to secure a stable income flow, hospitals need to secure a well maintained relationship between hospital and patient. For this reason, hospital management should emphasis the importance of customer orientation, by providing better health care service only then will it out perform competitions.

In addition, both internal marketing and innovative concept emphasized on building an effective inter-communication system, this encourages members from inter-functional departments in reaching a common ground. When members shared the same managing philosophy with administrators that would benefit future communication. This research discovered that, the performance of internal marketing will affect the level of innovation attitude thus it is recommended besides encouraging inter-functional cooperation and coordination through internal marketing, integrating and maximizing internal resources as well as cross functional communication in designing innovative health care service delivering diverse service.

REFERENCES

References available upon request from Yafang Tsai.

OPERATIONS MANAGEMENT AND INFORMATION SYSTEM NEEDS IN INDIAN SMALL & MEDIUM ENTERPRISES

Rajiv K. Srivastava, Indian Institute of Management Lucknow
Lucknow, UP 226013 India, 91-522-2736613, rks@iiml.ac.in
Rahul Pandey, IGSA Labs
Hyderabad, AP 500007 India, 91-9731738131, rahul.pandey@igsalabs.com

ABSTRACT

Several issues of developing country Small and Medium Enterprises (SMEs) are inadequately covered in management literature. This paper is an outcome of several years of empirical work done by the authors on small firms in selected regions of India as an attempt to address this gap. The objective of the studies has been to understand how Operations Management is practiced in the Indian SME sector, and how such firms' operational performance and competitiveness can be facilitated by proper design and use of Information Systems/ Information Technology (IS/IT) and Decision Support Systems (DSS). We have adopted a case-based approach with eleven in-depth studies of Indian SMEs supplemented by sustained interactions with several SME owner-managers over the years in various professional and technical forums. Based on our observations, we suggest a possible roadmap for Indian SMEs in their progression towards formal Operations Management systems facilitated by adoption of IT/IS and DSS.

Keywords: Asian Business, Small & Medium Enterprises, Operations Management, Management Information Systems, Decision Support Systems

INTRODUCTION AND STUDY METHODOLOGY

Small and Medium Enterprises (SMEs) play an important catalytic role as developing economies strive for widespread industrialization. The Indian top leadership has acknowledged that the country cannot depend on a few large industrial houses and capitalists for driving the industrialization process. The government has set ambitious targets for the sector's growth, and has also launched a number of federal and state level legislations and policy initiatives to facilitate its development.

There are over 15 million SME units in the country, with the second-largest total employment after agriculture. The sector contributes about 40% of the total industrial value addition, and about 35% of national exports [1]. About 400 modern industrial clusters, some of them world-class, have taken root over the years in products such as hosiery, knitwear, brass products, and gemstones. Apart from these, over 3000 traditional clusters such as handicrafts and handlooms also run mostly as SMEs and face similar management challenges. Such firms are also increasingly expected to integrate into global supply chains.

SMEs in developing countries face certain problems that are unheard of in advanced parts of the world. One, most medium sized firms of developing countries would be categorized as ‘small firms’ relative to medium firms of advanced economies. That is, when we compare firms across the globe, there are a large proportion of very small to small firms in developing countries. Two, small firms in developing countries face more restricted access to capital and infrastructure support than their counterparts in advanced countries. Three, due to poor access to capital, these firms also have poor access to skilled human and technological resources.

The current worldwide economic slowdown has significantly impacted the SME sector, reducing demand and hitting revenue realizations. It has also made access to external financing and resources even more difficult, and set back growth plans. It is therefore important, as well as an opportunity, to remove waste and streamline processes through the involvement of partners, suppliers and customers. Unfortunately, due to their inability to attract trained technical and management specialists as employees or consultants, they are unable to systematically pursue such initiatives.

Most empirical efforts in management literature, research, and teaching tend to be biased towards the large organized sector. At several institutional and professional forums over the last few years, there has been emphasis on the need to study the so-called “under-managed” sector, particularly small firms, traditional local industry, and the non-corporate sector. In order to motivate qualified professionals to create wealth and jobs by setting up their own ventures, several institutions have been setting up Incubators and Entrepreneurial cells. However, most of the SME specific research presently comes from the advanced countries. Although the organization, technology and management in large firms in the developing countries have caught up significantly with those in the advanced countries, the SMEs in the developing world are still quite different from their counterparts in the advanced markets. Several issues of developing country SMEs are still inadequately covered in Operations Management (OM) literature.

In light of these factors, we felt that it would be interesting to investigate how OM is practiced in the Indian SME sector, and how their owner-managers’ functioning can be facilitated by proper design and use of Information Systems/ Information Technology (IS/IT) and Decision Support Systems (DSS). This paper is an outcome of several years of empirical work done by the authors on small firms in selected regions of India. The objective of the study was to observe the status of such firms on various aspects of OM, understand their challenges, and to identify possible steps toward improvement of their operational performance and competitiveness. We have adopted a case-based approach with eleven in-depth studies of Indian SMEs owned mostly by ‘technopreneurs’ (entrepreneurs with prior technical education). These span a range of national regions, product types, technologies, sizes, and stages of life cycle evolution. The case studies have been supplemented by our sustained interactions with a number of SME owner-managers over the years in various professional and technical forums.

The firms that we studied were characterized by fairly centralized owner-control of business in general and operations in particular. While a few were set up as ancillaries or suppliers for large firms, some were independent but at times with limited and stable major customers. The nature of operations and degree of uncertainty varied with the position on the Independent – Ancillary continuum. The key operating priorities were typically low cost and flexibility, with some attention to capability and delivery especially for ancillaries and captives.

The functional OM aspects studied include the product/ market context and operating priorities, product & process design/technology, factory capacity and layout, production planning & scheduling, materials management, quality, maintenance, and sourcing & supplier management. We have also documented examples of technical and managerial innovations carried out over time, as well as the nature of record-keeping and use of IS/IT.

TECHNOLOGICAL INNOVATIONS

While there is little breakthrough technological innovation in the majority of small firms, several firms whose owners came from engineering background related to their business were distinctly more aware about modern technologies and technical practices in their manufacturing processes and products. Technical passion was one of likely inspirations for many such owners to start business in their area of expertise.

Such owners, by virtue of their inspiration, pursue small, incremental technical innovations on an ongoing basis. When resources permit, they invest in and upgrade key technologies. However, since such resource cushions are not easily available, they often engage themselves and their technical staff in exploring incremental, retrofit type innovations. We observed many examples of such indigenous, low cost innovations in the firms we visited - a result of the technical bent of owners and working capital constraints.

Some examples of incremental innovations we observed include adjusting orientation of one side of the lathe to suit turning operation on a particular component; welding stoppers on the lathe bench to prevent excess longitudinal movement (an example of pokayoke); analyzing different options for placement of risers and gates in the mould designs using a casting simulation software; use of air compressor to spray concrete mixture during a ferro-cement construction; devising a hanging mechanism in microscope for providing vibration insulation; and modifying the component designs to avoid usage of precision grounded screws.

OPERATIONS MANAGEMENT PRACTICES

Through the case studies we have observed that Operations practices in such SMEs tend to be more ‘engineering driven’ than ‘management system driven’. While they are reasonably open to fresh infusions of core technologies, the same level of initiative is not visible in financial, marketing and managerial innovations. Perceptions of productivity, quality and maintenance are linked to technology rather than operational systems. The majority of small firms do not have formal systems for OM functions and activities. Systematic capability assessment and development through process control/improvement and workers’ training/involvement are usually under-emphasized, despite problems of inventory costs, quality and lead-time as volume and variety increase over time. The overall effect is therefore that practices in these areas are not well organized.

SMEs which become reasonably successful encounter multiple strategic, operational and organizational challenges in managing their growth to the next level. The increasing variety and volume requirements put pressure on hitherto non-formal operations systems and structures. Such non-formal operations systems serve the purpose well in the start-up phase of a firm when the need is to make the operating facility and resources flexible in order to respond to needs of different small-order customers, while keeping overhead cost low. However, as a firm passes the first phase of survival and attempts to attract and retain some good customers, the thrust of deliverables shifts from 'only low cost' to a portfolio of performance measures like cost, quality and delivery. It is then that the non-formal systems become a barrier when most customers look for reliable performance before committing longer-term and larger orders. In other words, the owners of small firms now have to cope with increasing demands of key existing as well as new customers, which result in pressures to concurrently and consistently deliver on multiple performance dimensions.

Most firms suffer heavy costs due to undermining systematic OM approaches. For instance, a foundry's owner, a heat treatment unit, an auto-component supplier, and a pressure gauge manufacturer acknowledged steady deterioration in process performance over time after initial purchase of equipment. Not having any structured process control methods resulted in high incidence of defects and rework at advanced stages, leading to both high semi-finished inventory and long lead time. We also observed that in most cases Inventory costs are not properly estimated. Absence of formal forecasting, production planning, scheduling and materials management often results in excess inventory of raw material, work-in-process and finished goods, and long lead times. Workers are rarely involved in a systematic manner in improvement initiatives.

Owner-managers therefore need to conceptualize and manage the transition from non-formal to formal systems for coordination, control and improvement systems for their operations, with corresponding changes in structures. A good understanding of OM coupled with Information and Decision technologies can play a major role in driving such initiatives and enable innovation, adaptation and fast response. However, we have found that the owner-manager stays involved in routine operations rather than focusing on more important areas such as strategy/ planning for new markets, new product development, new materials, new vendors, new process technologies, capacity enhancement, etc.

INFORMATION SYSTEMS/TECHNOLOGY ADOPTION

With absence of formal systems, there is little IS/IT usage in operations and basic record keeping is mostly manual. Computerized transactions recording and use are confined to accounting and personnel information, and in some cases, customer/vendor order/invoice transactions. Relatively lower skill of employees in operations also implies lower awareness of computers and IS.

Basic level of IS adoption to support desired formalization of operational systems becomes advisable at this stage. This also facilitates management of this shift by relieving the owner-manager of day-to-day tasks. Since absence of systematic data recording/storage precludes design of formal operations management systems, it is crucial to begin with the former. Data recording/storage of relevant operations information/transactions like bill of materials, item classification, procurement, inventory storage (receipt/consumption), material flow, process activity/output and quality checks, can lay the foundation for designing useful rules for managing operations in these areas.

This initial phase of IS design & implementation also needs to emphasize the importance of data disciplining, and motivate/ facilitate the learning and use of simple OM techniques for better coordination, standardization and analysis. Basic computer training of selected personnel and use of IS for systematic recording/storage of relevant operations data is a natural initial step to toward deployment of IS/IT to support enhanced operations.

From a design perspective, IS and OM systems at this stage need to incorporate user-driven maintainability and upgradability of IS and OM systems. Such systems should preferably not require major capital investments and overheads, and therefore make use of inexpensive software and hardware, especially the deployment of accessible stand-alone systems through open source rather than proprietary packages.

ANALYTICAL MODELS FOR DECISION SUPPORT

Even when SMEs deploy contemporary technology, they usually tend to automate only their technical processes, data collection, and accounting systems, ignoring perhaps the most important – automatic decision making. Decisions tend to be concentrated in the hands of the owners or a select few among top management.

Most small firms do not use analytical models/tools for decision support. Some firms driven by technically qualified owners do use such models, but such applications are limited to designing of product and/or process. The managerial decision making is almost always based on judgment and experience of owner-managers.

Even though there is clear utility of simple analytical tools for decisions relating to materials management, operations planning and scheduling, and quality control, their use is obviated by the absence of both systematic operations data recording/storage and IS. A basic level of experience of both will facilitate use of simple analytical tools.

STRATEGY FOR TRANSITION TOWARD FORMAL OPERATIONS SYSTEMS

In the next phase, the IS capabilities need to be enhanced to include basic what-if analysis so as to facilitate both shift to formal OM systems and explicit support to operational, tactical and strategic decisions that fit with the changing market demands. This initiates the transition towards DSS by improving the quality of decision making, with gradual evolution from spreadsheet/ database tools to more capable quantitative models. Some examples of such systems supporting ‘infrastructural’ shifts include: Evaluation of different stocking/ ordering policies; Process quality control methods; Cost/ revenue implications of quality defects, delivery delays, etc. Likewise, some examples of systems supporting ‘structural’ shifts could be: Evaluation of different layouts; Capacity planning and evaluating investment options; Accessing information on new technologies, locations, suppliers, etc. and their evaluation.

Based on the stage at which a firm is in terms of the degree of formalization of operations and IS/IT systems, a logical progression towards formal and effective systems could be envisaged. The figure below indicates such a transition for a small firm that has almost no formal systems to begin with:

Definition	→	Measurement & Storage	→	Decision/Analysis Systems
Operations Data Identification; Item Listing/Coding; Bills of Materials; Item Classification; Process flow		System for data recording; Data measurement & recording (materials procurement, stocking, flow; Process status; etc); Data storage		Inventory control; Quality control; Capacity planning; Production planning & scheduling

While the range of capability development and improvements needed by Indian SMEs seems formidable, its achievability has been demonstrated successfully in the auto component and other such established sectors. For example, regular and rigorous system and process audits from key customers have contributed in inculcating discipline and knowledge. Such firms have also benefited from training and vendor development programs by larger customers as well as through associations and clusters. Many small firms have already made significant improvements, for example over 15000 units are now ISO9000/14001 certified. In fact as per the findings of a national Manufacturing Excellence Award program started five years back, the SME sector has shown comparable overall improvements vis-à-vis large industry, and some top small firms are almost on par with their larger counterparts [2]. However, much more concerted effort is needed to make such improvements systemic and pervasive.

REFERENCES

[1] Chavan, S. “Recessionary Wave Pounds SMEs”, Industry 2.0, December 2008, 38-39.

[2] Veeraraghavan, M. “SME Sector in Manufacturing”, Industry 2.0, March 2008, 40-43.

COLLECTING RESEARCH DATA IN ASIA USING SURVEYS

Thomas R. Tudor, University of Arkansas at Little Rock, (501) 569-8895, trtudor@ualr.edu

ABSTRACT

This paper examines various successful methods for collecting data for research in Asian countries. An emphasis is placed on survey data. Topics include the design of surveys for use outside of the United States, choosing translation services, testing surveys, picking survey collection methods, gathering data, and dealing with various cost factors. Extensive advice is also given on getting survey research completed in Asia approved through the Institutional Review Board located in all academic institutions. Examples are given using a recently completed survey in which data was collected in Taiwan and Thailand.

INTRODUCTION

Collecting data for research is often a difficult process. Using surveys as the main source of data is often the most difficult. While the use of on-line surveys has reduced the use of the classic survey through the mail method, obtaining a representative credible sample still remains a challenge. This challenge is magnified if data is needed outside your country of origin. In my situation as someone from the United States, I needed to collect data in Asia and eventually Europe. I chose Taiwan and Thailand as my first comparative countries because I wanted to start in Asia and I was very familiar with both these countries.

SURVEY DESIGN

My first step was to design a survey in English with no jargon that would translate well into other languages. Survey questions were developed through a review of other research, much brainstorming, and with advice of peers after many drafts. Because this survey would be used extensively for response comparison purposes from respondents in different countries, it was important that the questions chosen were polished and that no area was left out. Although adding or modifying a question later would not have been a gigantic blunder, the ideal would be to keep all surveys consistent. The list of questions was kept somewhat short with 24 multiple-choice questions chosen. Open ended questions were eliminated because of the additional time needed for respondents to fill out and because responses would have to be translated.

SURVEY APPROVAL

The next step was to have the survey reviewed by our Institutional Review Board (IRB). All research involving human subjects must be reviewed by IRB. This was a big challenge. Although my research fit the “exempt” category as far as federal regulations, the approval process is similar to "non-exempt" research requirements and this is true at most academic institutions. The first challenge was getting the English version of my survey approved. I did not want to submit the translated surveys with the English version until the English version was approved. The rational was that the translations would have to be completed again and this was an expense and extra time waster. However, certain IRB members would only informally approve the English version because the full IRB research proposal would not be officially reviewed until the entire packet was submitted. This entire packet would have to include the translated surveys as well along with many other administrative forms. I would recommend planning for at least three months of back and forth with IRB if research surveys will need to be translated and to deal with any unforeseen questions, concerns, or requirements. For instance, I had to also complete six on-line quizzed modules on proper human research practices.

SURVEY LEGAL AND ETHICAL REQUIREMENTS

The English version of my survey was eventually approved, but with many required changes, and with a requirement that specific wording be included in the instructions on taking the survey. This new wording added two paragraphs to the survey and included many disclaimers and legal wording. According to IRB, respondents must know that surveys taken without names are confidential and anonymous and that it is voluntary to take the survey. Respondents can also stop taking it at anytime. Respondents must also know that IRB had reviewed the survey and that it meets all ethical requirements under USA federal law and university policies. Contact information for IRB and the researcher must also be listed. Below is what was finally approved by IRB after multiple submissions:

This study is being conducted by Dr. Tom Tudor at the University of Arkansas at Little Rock in the United States. The below survey questions will help increase the understanding of sexual harassment perceptions in workplaces around the world. Your individual responses to this survey are confidential and anonymous because no personal identifiable data will be collected and no one other than this researcher will have access to your individual answers. This survey is also completely voluntary and you may stop taking it at anytime. If you decide to complete it, the survey should take no longer than 3 minutes. The return of this survey will serve as your consent to participate in it.

This study has been reviewed and approved by the University of Arkansas at Little Rock's Institutional Review Board (IRB). The IRB has determined that this study meets the ethical obligations required by USA federal law and University Policies. If you have any questions about this study, you may contact the researcher at trtudor@ualr.edu or 011-501-569-3353. If you have any questions regarding your rights as a research subject, please contact Michelle Morgan at 011- 501- 569-8895 [1].

SURVEY TRANSLATION

The next step was to have the English version of my survey translated into Mandarin Traditional Chinese for Taiwan and Thai for Thailand. The use of my wife to translate for the Taiwan survey was rejected unless I also had someone else translate the translated survey back into English to make sure it translated correctly. IRB was having a difficult time approving a translated survey that they could not read. They would not hire their own translator to verify the work of my wife and that of another to translate the survey into Thai for a second check on the original translation. This dilemma was going to add to research expenses and push forward the start date significantly. This became a panic situation

Fortunately, I came across a translation firm that actually reduced my translation expensive and performed the work even faster than I had originally planned. SDL is the largest and most respected translation firms in the world. Sample clients have included Adobe, Sony, Microsoft, and many federal agencies [2]. They reasonably offer a double translation service in which one translation is completed and a second translator reviews the work of the first with both being native language speakers. SDL also certifies their translations as accurate and gives an official certification. Cost is based on word count and the English version of my survey used 880 words for the twenty-four questions and survey instructions. The Thai translation cost $88.00 and took less than a week for SDL to complete the work. The Mandarin Chinese translation was more expensive because Taiwan residence read traditional Mandarin Chinese characters while mainland China residents read simplified Mandarin Chinese characters. The cost was increased to $178.00 but would have been $88.00 if translated into simplified Mandarin Chinese characters to use in mainland China. I have to assume the labor is more expensive in Taiwan because SDL offices in each native country do the translation, however, it would have been a major blunder if I had the survey translated into the wrong Chinese characters.

Because of the professionalism and reputation of SDL, IRB accepted their translations as accurate. As my own further check on the translation work, I had international friends read the translated surveys to make sure they matched the English version. They did.

TRANSPORTING SURVEYS TO ASIA

IRB required that one original copy of each translated survey be stamped by IRB. I made 250 copies of each survey (Taiwan and Thailand versions) and carried them with me to these countries. I was told that I could have had them copied more cheaply in Asia but I was not positive of this fact and chose to have them copied here. I also had no way of getting reimbursed there either for my copying expenses. The surveys did add weight to my luggage. They made it safely though with no damage. I was ready to start distributing them.

SURVEY TO RESPONDENTS

I chose to give out my surveys to employees located around major corporations in both Taiwan and Thailand. This was time-consuming but the response rate was very high and surveys were typically fully completed. I have some company contacts now but I lacked them on my initial trip. If needed in the future, I feel I can get some Asian employers to have their employees do them for me. For IRB approval, I could not just go and find employers after arriving in these countries. I had to actually list them and get written permission. This was impossible for the timeframe I had to do this work and get IRB approval. IRB would have also had to approve employers if I did an internet survey. There are some firms that pay for employees to do surveys on-line in countries all over the world and researchers pay the firms. I wasn't comfortable with this method but it could be a good one if response rate is high and answers are credible. I feared a lack of credibility and could not find another researcher that had used this method.

CONCLUSION

Having respondents complete your research survey in countries not of your origin or language is a challenge. There is definitely a learning curve in the design of survey questions, getting questions translated, and having respondents complete them. In addition, working with IRB can be time-consuming, frustrating, and also rewarding. I initially thought it was ridiculous that exempt research had to jump through so many IRB hoops. However, IRB did ultimately help me create better surveys and forced me to think much further ahead as to the planning of my research. My research quality was better with their help and I am grateful. However, working with IRB can be a long process and their full approval did not take place until less than a week before I was to fly to Taiwan as my first destination before flying to Thailand. My advice is to get IRB involved months in advance if possible. This is particularly true if you have never worked with IRB.

REFERENCES

[1] Tudor, T. R. Sexual Harassment Perception Survey (unpublished instructions), 2009.

[2] SDL Inc. at http://www.click2translate.com/

SPATIAL ANALYSIS OF INVESTMENT LOCATIONS. DOES COUNTRY BORDERS MATTER ? THE CASE OF JAPANESE MULTINATIONAL COMPANIES IN EU

Martins Priede, Yokohama National University,
Tel/Fax (+81) 45339-3659, d07lb907@ynu.ac.jp

ABSTRACT

This paper empirically analyzes the location choice of Japanese multinational companies in EU. By using sample of 1023 Japanese investments, conditional logit and two level nested logit models are estimated in 18 country and 236 regional locations over the period from 1995 to 2004. Regions have been grouped and estimated in three different structures. The results of nested logit when regions are grouped in their respective country nests show that investors don't select particular country before choosing the region in it, but select wider geographical area and then chooses region within it, without significantly influence by existing country borders.

Foreign direct investment, Japanese multinationals, nested logit, European Union

1. Introduction

Nowadays companies are free to choose locations, which best suit their interests. Although some industries like mining and retailers are limited to certain locations where natural resources or customers are, most manufacturing and wholesale industries are free to choose where to be located. Moreover, transportation costs have been decreasing and modernization of telecommunications have facilitated and fastened flow of information. For instance, in EU regional integration has greatly reduced trade costs and difficulties to serve customers in different country or region in Europe. [2]

On the other hand, companies are exposed to greater competition in such an integrated area, because the number of potential competitors increases as well. As a result, one of engines to stay competitive in such wide and integrated markets is to select the location, which best suits their strategic, operational and financial interests.

In previous studies, several factors emerge according to which multinational companies choose investment location. The first one is market – companies want to serve markets where demand is high (market potential factor and companies can enjoy economies of scale), the second one is costs – companies want to locate where operational costs are smaller (labor costs, expenses of renting/buying property, taxes etc.), the third one – companies tend to choose the same region as companies with similar attributes are located – related industries and same country of origin (agglomeration effect) and the forth one – resource availability as companies don't want to face shortages of labor force, government services, natural resources etc.

By using nested logit model, the paper analyzes location choice of 1023 Japanese investments in 236 regional locations belonging to 18 European countries over the period from 1995 to 2004.

Proposed nested logit has been used in several papers dealing with location choice in Europe. Head [5] analyzes location choices of Japanese investment in nine EU countries. This study explores various market potential indices, notably the index developed by Harris [4] and Krugman [8]. It has been found that market potential is a significant and positive factor for investors' location choice.

Earlier paper by the same author [6] uses conditional logit to analyze location choice of 751 Japanese manufacturing plants in US in the 1980s. They found strong agglomeration effects on location choice, especially for companies within a similar industry or belonging to keiretsu. The authors try to different groups of countries e.g. North-South, Anglo-Continent-South etc. It has been found that country boundaries don't matter in the case of EU and US multinational companies.

By using conditional logit model, Alegira [1] analyzes location decisions of European firms across Europe with large sample size of 4 803 foreign investment projects in 246 regions. This study considerably extends geographical areas of previous location choice papers. Similarly to previous research, market potential has been found to have positive and significant influence.

Heterogeneity of investors has been analyzed with the mixed logit method in Rasciute [10] paper on the location choice of foreign investors in 13 Central and Eastern European Countries. This study reports high heterogeneity of investment location decisions. Market effect also has been observed, suggesting that larger host country will be more likely to be selected and this effect is stronger for larger firms.

There are also studies comparing investment location choices in different areas of Europe. Disdier [3] analyzes location choice of French firms in Eastern and Western Europe by using both conditional and nested logit methods. Results suggest that French firms will choose countries with already established French firms of the same industry.

Several papers include unemployment rate, but results are mixed for this factor. As Disdier [3] points out that high unemployment rate might suggest labor market imperfections, but on the other hand from investor's view point might signal availability of large labor supply.

This research contributes previous studies in such a way as it extends geographical spectrum of analysis, lowers statistical level of regional analysis and introduce several new nested logit structures.

2. Theoretical models

Discrete choice models are widely used to analyze firms' location choice. Most accepted model in recent studies is conditional logit model developed by McFadden [9] for the analysis of consumers' utility, but this model can be well applied to analyses of investors' location choice. The proposed model maximizes firms' profits – Japanese investors will choose particular location if that location offers highest profits among other possible alternatives.

As rational decision makers, firms choose to invest in location, which offers biggest profit π or profit at specific location $\pi_j = V_j + \varepsilon_j$ where V_j is attributes to location j and ε_j is error term. Here, subscript $j = 1, \ldots, J, [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_J]$.

Linear expression is $V_j = \beta X_j$ where X_j is a vector of observable characteristics of location j and β is a vector of estimated parameters. Coefficients β are further estimated by maximum likelihood technique. Equation (1) below gives the probability of a firm choosing location j over location k.

$$P_j \equiv prob(\pi_j > \pi_k) = prob(\varepsilon_k < \varepsilon_j + b(X_j - X_k)), \forall k \neq j \qquad (1)$$

Choosing j over i depends only on two alternatives, only then the independence of irrelevant alternatives (IIA) condition holds. Nested logit allows variance to differ across the groups while maintaining the IIA assumption within the groups in this case countries and geographical areas, but IIA will not hold between groups.

A firm earns profit $\pi_{ij} = V_{ij} + \varepsilon_{ij}$ at region i in country j, where $V_{ij} = bX_{ij} + aY_i$. Equation (2) gives probability of nested structure logit.

$$P_i = \frac{e^{\alpha Y_i + \sigma_i I_i}}{\sum_{m=1}^{I} e^{\alpha Y_m + \sigma_m I_m}} \qquad (2)$$

Here in Equation (2), I_i is the inclusive value or τ, which measures dissimilarity of the nest. When the coefficient τ takes values 1, nested logit turns into conditional logit model and there is no tree structure observed. When estimated coefficient τ takes value 0, choices in group are considered to be perfect substitutes. [3]

For the purpose of this paper the following three non-overlapping nested structures of regions were created, regions nested in their country borders, regions of South-East and North-West Europe and wider area geographical nests.

3. Data and sources

Location data about investment of Japanese companies were collected from *Toyo Keizai* database of year 2006, Eurostat and KPMG survey [7] [11]. Due to limited data availability of statistical data about Denmark, Romania and Bulgaria, data of those countries were removed from the dataset. Nests, which contain only one alternative, the dissimilarity parameter could not be identified. Therefore, due to only one national subdivision of statistical NUTS2 level, countries such as: Cyprus, Estonia, Lithuania, Latvia and Malta were further removed from the dataset. In addition, overseas regions and dependent territories were excluded from the dataset.

Ten years time period were further split into three non-overlapping time periods to observe factors, which determine location decision over time. First period is from year 1995 to 1997, second period - from 1998 to 2000 and third period from year 2001 to 2004.

On the left side of the logit equation, the dependent variable is qualitative variable indicating whether multinational company chooses particular region. Variable takes value 1 if Japanese investor(s) chooses the region in the particular year, otherwise value is 0.

On the right side, independent variables are region and country specific. Market size is characterized with regional GDP and GDP per capita of particular region, both expressed in euro currency for all countries. GDP per capita might characterize income as well as labor cost level. High income levels might be attractive for Japanese investors as it might indicate high labor quality, but in the same time might have negative influence.

Cost of location is captured by such variables as costs of owning and renting real estate property and corporate tax rate. Included, corporate taxes have negative influence on investors' interest in the country and region, as higher taxes reduces net profits.

Agglomeration variable is the number of Japanese investments in region. Japanese investors are more likely to select region where is already established companies, because of positive spill over of information and possible cost savings in procurement. Infrastructure in a region is measured by road density, which is calculated as the ratio of length of road network per regional area.

4. Results and conclusions

Conditional logit model shows that such factors as regional income level, infrastructure advancement level and corporate tax rate is not significant for Japanese investors' location choice. On the other hand, significant factors are regional GDP, number of other Japanese investment projects in the region, expenses on renting/buying property and unemployment rate. When time period is divided in three sub-periods, analysis reveals that income level, expenses on property, unemployment rate has become significant during sample period. In addition, analysis suggests that during observed period Japanese investors have becomes more cost conscious and prefer regions with cheaper labor and property costs.

Further analysis of nested logit model improves results of conditional logit model. Although, two level nested logit, which divides investors' choice between South-East and North-East European regions, did not improve conditional logit model. When regions are grouped in their country nests, significance of model factors decreases, but inclusive value suggests that grouping regions by in their respective country groups are not valid and there are many cross-border influences. Region intra-dependence increases considerably during last sub-period.

Finally, when grouping regions in larger geographic area groups, which reflect their cultural and geographical linkages, significance of factors improves considerably. Inclusive value suggests all IIA are within selected groups. This suggests that Japanese investors will select larger geographic area and then region within that area. In this study, these observed and important factors are regional GDP, expenses on property, number of other Japanese investment projects and unemployment rate. Factors, which enhance region's chances of being selected, are increase of regional GDP and number of other Japanese investment projects in the region. Factors, which decrease region changes of being selected, are increase of property costs and unemployment rate.

During sample period, it has been observed that there is no significant border effect in Japanese investors' approaches to EU. Yet, due to increased competition, Japanese investors tend to select regions with lower operational costs, which allow them to be more competitive. This tendency especially is observed in the last period in considerably enlarged EU.

5. References

[1] Algeria, Rodrigo “Countries, Regions and Multinational Firms: Location Determinants in the European Union”, London School of Economics, 2006, viewed 27 July, 2009, http://www-sre.wu-wien.ac.at/ersa/ersaconfs/ersa06/papers/143.pdf

[2] Clark, C. F.Wilson and J. Bradely “Industrial location and economic potential in Western Europe”, *Regional Studies*, 1969, Vol. 3, pp. 197-212

[3] Disdier, Anne-Celia and Thierry Mayer "How Different is Eastern Europe ? Structure and Determinants of Location Choices by French Firms in Eastern and Western Europe", *Journal of Comparable Economics*, 2004, Vol. 32, pp. 280-296

[4] Harris, Chauncy D. "The Market as a Factor in the Localization of Industry in the United States", *Annals of the Association of American Geographers*, 1954, Vol. 44 (4), pp. 315-348

[5] Head, Keith and Thierry Mayer “Market Potential and the Location of Japanese Investment in the European Union”, *The Review of Economics and Statistics*, 2004, November 2004, 86(4): 959-972

[6] Head, Keith, John Ries and Deborah Swenson "Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investment in the United States", *Journal of International Economics*, 1995, Vol. 38, pp. 223-247

[7] KPMG International *KPMG’s Corporate and Indirect Tax Rate Survey 2008*, 2008, viewed 27 July, 2009, https://www.kpmg.es/download/TaxRateSurvey08.pdf

[8] Krugman, Paul "A Dynamic Spatial Model", *National Bureau of Economic Research*, 1992, viewed 27 July, 2009, http://www.nber.org/papers/w4219.pdf

[9] McFadden, Daniel “Econometric Analysis of Qualitative Response Models”, editor: Griliches, In *Handbook of Econometrics*, 1984, Vol.2, Elsevier/North-Holland, Amsterdam

[10] Rasciute, Simona and Eric J. Pentecost “The Location of Foreign Direct Investment in the Central and Eastern European Countries: A Mixed Logit and Multilevel Data Approach", 2007, *Discussion paper series of Loughborough University*, viewed 27 July, 2009, http://www.lboro.ac.uk/departments/ec/RePEc/lbo/lbowps/Nested_Logit.pdf

[11] *Toyo Keizai* database *Kaigai Shinshutsu Kigyou 2006*, 2006, CD-ROM database

A COMPARISON OF THREE CASE STUDIES INVOLVING THREE STAKEHOLDERS, EACH WITH UNIQUE OBJECTIVES, WHO HAVE ENTERED INTO SURROGACY AGREEMENTS

Cynthia L. Knott Ph.D.
Marymount University, 2807 N. Glebe Road, Arlington, VA 22207
cynthia.knott@marymount.edu
703-284-5727

Susanne Ninassi, J.D.
Marymount University, 2807 N. Glebe Road, Arlington, VA 22207
susanne.ninassi@marymount.edu
703-284-5934

Alyson Eisenhardt, DHSc
Marymount University, 2807 N. Glebe Road, Arlington, VA 22207
alyson.eisenhardt@marymount.edu
703-284-4984

Northeast Decision Science Institute

March 2010

Abstract

This paper builds on the previous work of Knott, Ninassi and Eisenhardt's 2009 paper, "Using the Analytic Hierarchy Process to look at the Tradeoffs among Stakeholders in the Case of Surrogacy, When the Initial Intention of the Parties Involved Change after the Pregnancy Has Ensued." This paper proposes a new methodology for analyzing cases involving stakeholders in disputed surrogacy arrangement.

The methodology utilized is the Analytic Hierarchy Process (AHP), which allows users to weigh different criteria based on their own opinions and background. Three case studies involved a surrogacy arrangement among three stakeholders: the biological mother, the biological father, and the surrogate mother. The base for the model was that all parties had entered into an agreement, and that after the agreement had been made, a situation changed, such that a decision as to who had parental rights over the fetus became an issue. The criteria used in the decision making process in the three case studies were legal, health and ethical. The criteria differed based upon each stakeholder's point of view.

In this paper, the three models were developed using the AHP methodology and judgments were gathered for each of the three cases to derive the priorities for the criteria in each model. Each author took the role of one of the three stakeholders.

Legal and Case Studies Background

As previously discussed in the prior paper (Knott, Ninassi, Eisenhardt 2009), the legal background for surrogacy cases follows.

Historically, when the custody of a child is contested, the courts will look to the standard of what is in the "best interest of the child" in reaching its decision.* However, with the advancement is reproductive technology, such as in vitro fertilization, the courts must grapple with the issue of who is the legal parent. Three California state cases are discussed to illustrate how courts are dealing with this issue.

The first case is *Johnson v. Calvert*, 851 P.2d 776 (Cal. 1993), cert. dismissed, 510 U.S. 938 (1993). This is case, the Calverts, a married couple, entered into a gestational surrogacy arrangement with Ms. Anna Johnson. Under this type of surrogacy arrangement, the wife's egg inseminated with the husband's sperm was transferred into the uterus of another woman, the surrogate, for gestation and birth. As part of the contract, Ms. Johnson agreed to relinquish all parental rights. During the pregnancy, the relationship between the Ms. Johnson and the Culverts deteriorated. The Calverts sued for a declaration that they were the legal parents of the unborn child. Ms. Johnson sued to be declared the legal mother.

The California Supreme Court held that although the state statute "recognizes both genetic consanguinity and giving birth as means of establishing a mother-child relation,

*Hames, J. & Ekern, Y. (2006). *Introduction to Law (3rd ed.).* New Jersey: Pearson Prentice Hall when the two means do not coincide in one woman, she who intended to procreate the child-that is, she who intended to bring about the birth of a child that she intended to raise as her own-is the natural mother." *Id.* at 782. The court held that the gestational mother had no parental rights, and the agreement was not inconsistent with public policy nor was the termination of the surrogate's claims to the child was not unconstitutional.

The next case is *In re Marriage of Moschetta*, 30 Cal. Rptr. 2d (Ct. App. 1994). In this case, the husband and wife, Robert and Cynthia Moschetta, entered into a traditional surrogacy arrangement in which Ms. Elvira Johnson, the surrogate, was impregnated with Mr. Moschetta sperm with the prior understanding that the resulting child would legally be the child of Mr. Moschetta. Ms. Johnson also agreed to terminate her parental rights, and Mrs. Moschetta would then adopt the child. After the baby was born, the marriage deteriorated, and Mr. Moschetta filed for divorce. The court was asked to determine the parental rights of the wife and the surrogate. The trial court held that Mr. Moschetta and Ms. Johnson were the legal parents of the child and should have joint custody. Mr. Moschetta appealed stating that contends that his wife is the legal mother.

The California Court of Appeals held that according to the state statute, the surrogate is the legal and natural mother genetically and by giving birth. The court noted that this case differed from the *Johnson* case since the wife, Ms. Moschetta, was not the child's genetic or biological mother. The court reasoned that since the state statue was clear there was no further need to look to the ruling in *Johnson*, and therefore held that the surrogate, Ms. Johnson, was the legal mother.

The third case is *In re Marriage of Buzzanca*, 72 Cal. Rptr. 2d 280 (Ct. App. 1998), reviewed denied, (June 10, 1998). In this case, the Buzzancas, a married couple, agreed to have an embryo genetically unrelated to them implanted in a woman, the surrogate, who would carry and give birth to the child for the couple. Before the birth, the Buzzancas split up, and the question before the trial court was who were the lawful parents. The trial court determined that the child had no lawful parents. The parties appealed.

The California Court of Appeals held that Mr. Buzzanca was the lawful father because "…there are times when *fatherhood* can be established by conduct apart from giving birth or being genetically related to a child'. *Id.* at 282. The court relied on the finding in *Johnson* and found that Ms. Buzzanca was the lawful mother. The court held that

> …just as a husband is deemed to be the lawful father unrelated to him when his wife gives birth after artificial insemination, so should a husband *and* wife be deemed the lawful parent of a child after a surrogate bears a biologically unrelated child a their behalf. "In each instant, a child is procreated because of a medical procedure was initiated and consented to by intended parents. *Id.*

Summary of the Three California Cases

In the three California cases, the courts first looked to the California state statue to help determine the legal status of parties. The courts also looked at the intention of the parties at the time they entered into a surrogacy arrangement to help determine the legal parent(s) of the child.

In *Johnson*, although the surrogate was the birth mother, she had no genetic connection to the child. Relying on the statue and the intentions of the parties, the California Supreme court held that the husband and wife were the legal parents.

In *Moschetta*, the surrogate was the birth mother and was genetically connected to the child. The court, relying on the statute, held that the husband and the surrogate were the legal parents. The court did not follow *Johnson*, since the surrogate in this case had the genetic connection.

In *In re Marriage of Buzzanca*, a genetic connection did not exist for any of the parties to the surrogate arrangement. The court, relying on *Johnson*, looked at the intent of the parties and held that the husband and wife were the lawful parents.

Application of Existing Law to Hypothetical Cases

Hypothetical Case 1

The Facts:

Husband and wife cannot conceive. They enter into a traditional surrogacy arrangement with a woman who has agreed to be artificially inseminated with the husband's sperm and to terminate her parental rights upon delivery. However, after delivery, the surrogate decides that she wants to keep the child.

Analysis based upon the three legal cases:

> Based upon the existing case law, it would appear that the *Moschetta* case would be followed since the facts in the hypothetical and the actual case are the same. The surrogate in the birth mother and is genetically connected to the child.

Hypothetical Case 2

The Facts:

Husband and wife cannot conceive. They enter into a gestational surrogacy arrangement where the wife's egg, inseminated with the husband's sperm, was transferred into the uterus of another woman, the surrogate, who agreed to terminate her parental rights. When the surrogate is seven months pregnant, the marital relations between the husband and wife deteriorates. The wife decides she does not want the child and will terminate her parental rights. She files for divorce. The husband still wants the child; however, upon hearing about the divorce, the surrogate decides not to terminate her parental rights.

Analysis based upon the three legal cases:

> Based upon the existing case law, it would appear that the *Johnson* case might be followed since facts in hypothetical and in *Johnson* are similar since the surrogate had no genetic connection to the child. Therefore, applying the case law to this hypothetical, the legal parent would be the father.

Hypothetical Case 3

The Facts:

Husband and wife cannot conceive. They enter into a traditional surrogacy arrangement with a woman who has agreed to be artificially inseminated with the husband's sperm. The surrogate carries the baby for two months, during a routine examination it is determined that the fetus has a medical issue. The surrogate wants to abort the fetus.

Analysis based upon the three legal cases:

> The California courts have not addressed this issue. However, the California Court of Appeals discussed that enforcing a surrogate contract could lead to many legal problems and questioned:
>
> > What if a surrogate mother took drugs or alcohol during her pregnancy in violation of her contract? Or wanted and abortion? Could the contract be enforced by court order and subsequent contempt? Would there be a "surrogate mother's tank" in the local jail. *In re Marriage of Moschetta*, 30 Cal. Rptr. 2d 893, 903 & n.23 (Ct. App.1994).

The Three Hypothetical Case Study Models with Weights and Outcomes

The three cases were developed into the Analytical Hierarchy Process (AHP) Models, which is a methodology developed by Thomas Saaty (Saaty 1994). The process allows users to enter judgments via a pair-wise comparison mode to derive the relative priorities of the criteria and the alternatives. The criteria that were considered in the hypothetical cases included; ethical, health and legal. The alternatives considered were: parental rights, custody rights, carrying to term and abortion. After the judgments were entered, the model was synthesized to determine the overall best alternative for each individual case. Depending on the background of the case, the priorities changed due to different scenarios. The findings are listed in the three tables below.

Hypothetical Case 1 Model

Criteria	Overall Criteria Priority	Surrogate legal parent	Father legal parent	Both legal parents
Health Issues	.069	.111	.111	.778
Legal Issues	.348	.111	.111	.778
Ethical Issues	.582	.107	.107	.776
Overall Alternative Priority		.109	.115	.777

Since there were no real health issues, except for those associated with a normal pregnancy, this criterion weighted low in the overall model. However, the fact that the health and the overall well-being of the child is best served by having both parents, this criterion was factored in the overall decision. In terms of the alternatives, both the husband (the father) and the surrogate are the biological parents. Therefore, both have parental rights over the child based upon the current case law. When weighing in the ethical factor, sharing parental rights was the overall best decision. Note, in this hypothetical, abortion was not an option since the biological parents wanted to keep the child.

Hypothetical Case 2 Model

Criteria	Overall Criteria Priority	Surrogate legal parent	Father legal parent	Both legal parents
Health Issues	.062	.319	.221	.460
Legal Issues	.653	.076	.766	.158
Ethical Issues	.285	.349	.168	.484
Overall Alternative Priority		.199	.496	.306

As in the case above, there are no health issues, and therefore that criterion weighted low. However, the legal criterion is more important than the ethical. In this case, the husband and the wife are the biological parents. Yet, the wife terminated her rights and now the surrogate, who has no genetic connection to the child, seeks parental rights. Although allowing parental rights for the surrogate was weighted high in terms of ethics, because of the high weight of the legal issues, the overall decision would be for only the father to have parental rights. Note, in this case, abortion was not an option because both the father and the surrogate wanted to keep the child.

Hypothetical Case 3 Model

Criteria	Overall Criteria Priority	Abort	Carry to Term
Health Issues	.088	.875	.125
Legal Issues	.243	.857	.143
Ethical Issues	.669	.667	.333
Overall Alternative Priority		.721	.279

In this case, the ethical issues weighted the highest, with legal the next highest, because the husband and wife still want the child even though the child will have a medical issue. Although the health criterion should, at first glance, be a priority, it weighted low. This resulted because of the ethical decision by the couple to keep the child even though the child will have a medical conditions. However, in the overall decision, abortion came out to be the best alternative. Even with respect to the ethical and legal issues, it still weighted high. In fact, abortion was the best alternative with respect to each criterion.

Conclusions and Future Research

The Analytic Hierarchy Process, as developed by Saaty (1994), is a methodology used to assist in high-level decision-making. AHP allows decision-makers to weigh criteria based on their own opinions and thoughts. This paper employed the usage of AHP methodology to analyze three related hypothetical cases involving parental rights and surrogacy. Judgments for each stakeholder were made based on the following components: legal, ethical, and health. The judgments were then assessed and evaluated to determine which criterion weighed highest for each stakeholder.

The researchers would like to explore the assistance of AHP models using these and other case studies in the classroom to discuss the tradeoffs with students when making decisions. We would also like to find a current case in litigation to see if our model comes to the same conclusions as the court. However, this presents challenges. As indicated in the current case law, the court looks at the biological connection of the parties to the child and then will examine the intent of the parties in determining parental rights.

References

Expert Choice Software ©. (2000)

In re Marriage of Buzzanca, 72 Cal. Rptr. 2d 280 (Ct. App. 1998)

Johnson v. Calvert, 851 P.2d 776 (Cal. 1993)

In re Marriage of Moschetta, 30 Cal. Rptr. 2d (Ct. App. 1994)

Knott, Cynthia L., Ninassi, Susanne & Eisenhardt, Alyson (2009) *Using the Analytic Hierarchy Process to Look at the Tradeoffs in the Case of Surrogacy When the Initial Intention of the Parties Involved Change After the Pregnancy Has Ensued*. In the proceedings of the 39th Annual Meeting of the Southeast Decision Science Institute, Charleston, South Carolina, February 18-20, 2009.

Saaty, Thomas L. (1994) *Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process,* Volume I-VI of the Analytic Hierarchy Process Series, Pittsburgh, PA: RWS Publications.

Planned Parenthood Affiliates of California Inc. (2004). Health Center. *Abortion at the First Trimester.* Retrieved from http://www.ppacca.org

A Quantitative Approach to Corporate Philanthropic Selection

Fariborz Partovi, Drexel University. (215)895-6611, Partovi@drexel.edu

1. Introduction

Philanthropy has been defined as an altruistic behavior, usually in the form of charitable actions or gifts, towards others in society [1]. Corporate philanthropy is often mentioned as one of the four social responsibilities of a company in corporate social responsibility (CSR) literature [2] and [3]. Philanthropy along with other components of CSR; economic, legal, and ethical [3] represents a high-profile concept that has strategic importance for many companies. Traditionally, corporations were restricted from contributing to social affairs; a 1954 Supreme Court ruling, the Smith Manufacturing Company versus Barlow case, involving a stockholder suit against Smith manufacturing for a contribution made to Princeton University removed this restriction [4]. The court established the "business judgment rule" that organizations were able to make contributions that in their judgments would promote the company's best interest. Since companies have given large sums to charities, oftentimes without understanding the interrelationship between corporate donation strategy and philanthropic initiatives. The increasing volume of requests for money, the corporate social responsibility movement, the responsibility of corporate management to make profit and the interrelationship between corporate strategy and CSR contribute to more accountability, and produce the demand for justification of process and the formalization of philanthropy selection. Traditionally, the process of selecting philanthropy projects is ad hoc decisions, at best based on a decision maker's experience, feelings and politics, essentially a nonscientific approach. Ad hoc decisions, if sometimes effective, do not consider all relevant factors and alternatives. Organizations in the United States are developing a more scientific approach to corporate social responsibility as they attempt to rationalize the donation decision-making process by institutionalizing formal structures for determining which projects to support. These include establishing a donation committee, installing formal policies and procedures that must be followed, carefully screening applicants, and using scientific techniques for selecting beneficiaries. This paper takes one company as a case study in developing formalization of philanthropy selection.

The author has made several modifications to ensure the confidentiality of the company and the readability of this paper. The subject company's name has been changed to ensure anonymity. The numerical data has been altered to protect the company's proprietary information. Finally, the particulars of the case have been condensed for readability.

2. Company Background

ABC is the world's leading manufacturer of all digital mass measurement weighted products for industrial use. It provides weight-feeding equipment used in the processed food, plastics and chemical industries. The company stresses operational innovation and continuous improvement. The company's headquarters is located in United States with various offices and plants all over the world. This case, undertaken in 2008, relates to a division of this company located in Maryland.

3. Data Envelopment Analysis

Charles et al [6] introduced DEA in 1978. DEA has since been successfully employed for assessing the relative performance of a set of firms, usually called the DMU, which use a variety of identical inputs to produce a variety of identical outputs. Assume that the there are *N,* DMUs producing *s* outputs using *m* inputs. Let the *e* th DMU produce outputs y_{re}, r =1,2,…..s using $x_{ie},$ i=1,2,…..m as inputs. The resulting output–input structure of DMUs is shown below. The objective of the DEA is to identify the DMU that produces the largest amounts of outputs by consuming the least amounts of inputs. This DMU (or DMUs) is considered to have an efficiency score equal to one. The efficiencies of other inefficient DMUs are obtained relative to the efficient DMUs, and are assigned efficiency scores between zero and one. The efficiency scores are computed using mathematical programming.

As DEA is now a widely recognized technique, it is not described in this paper. Interested readers are referred to [7], [8], [9] and [10]. There are different basic DEA models (See for example [7]).

One model for calculating the efficiency of a DMU is shown below:

$$\text{Max } E_e = \frac{\sum_{r=1}^{s} u_{re} y_{re}}{\sum_{i=1}^{m} v_{ie} x_{ie}}$$

Subject to:

$$\frac{\sum_{r=1}^{s} u_{re} y_{rj}}{\sum_{i=1}^{m} v_{ie} x_{ij}} \leq 1, \quad j=1,2,3\ldots\ldots n$$

$$u_{re},\ v_{ie} \geq 0$$

$$r=1,2,3\ldots\ldots s,\ \ i=1,2,3\ldots\ldots m$$

where the subscript i stands for inputs, r stands for outputs, and j stands for the DMUs. The variables u_{re} and v_{ie} are the multipliers to be determined by the above mathematical program, and the second subscript e indicates the eth DMU, and E_e is the efficiency of the eth DMU. The above fractional program can be easily solved by setting the denominator of the ratio equal to unity; one can obtain the following *output maximization* linear programming problem.

$$\text{Max } E_e = \sum_{r=1}^{s} u_{re} y_{re}$$

Subject to:

$$\sum_{r=1}^{s} u_{re} y_{rj} - \sum_{i=1}^{m} v_{ie} x_{ij} \leq 0, \quad j=1,2,3\ldots\ldots n$$

$$\sum_{i=1}^{m} v_{ie} x_{ie} = 1$$

$$u_{re}, v_{ie} \geq 0$$

$$r=1,2,3\ldots\ldots s, \quad i=1,2,3\ldots\ldots m$$

In a slight variation of this model, the inputs for all DMUs are equal and a constant. Here the above model will be reduced to:

$$\text{Max } E_e = \sum_{r=1}^{s} u_{re} y_{re}$$

Subject to:

$$\sum_{r=1}^{s} u_{re} y_{rj} \leq 1, \quad j=1,2,3\ldots\ldots n$$

$$u_{re} \geq 0$$

$$r=1,2,3\ldots\ldots s$$

DEA divides the units into two groups: efficient and inefficient; in practice, there is often a need to fully rank the DMUs, and identify best-performing overall DMU among many efficient ones. Many authors have proposed various ranking methods in DEA and the most popular ones are: (1) cross-efficiency model [11], (2) super efficiency model [12], (3) benchmarking [13], and virtual method [14]. Each of these ranking methods is next briefly explained. The reader interested in the details of these approaches is referred to [11], [12], [13], and [14].

4. The Model and Conclusion

This section implements the proposed mathematical framework based on rankings of local high school needs using DEA. In the past, although ABC Corporation was involved with county high schools, their efforts were diverse and were not strategic. By focusing on one school with the most need and most relevant to employees, ABC Corporation realized the benefits of benevolent philanthropy, maximally supporting community, employees, and the interest of corporation. Further, partnering with only one school consolidate resources and create efficiencies with the volunteer program and with the overhead department paperwork associated with ABC charitable donations. In 2008 ABC Human Resources, aided by a consultant, adopted the focused strategy. To ensure the fairness and applicability of the method, and limit subjective decisions, the department determined several objective criteria. After elaborated discussions among HR personnel assigned to the task, they determined under executive approval the two most important criteria for selecting a school: the number of employee's children who attend the school; and need as assessed by the school's financial need, attending students' financial need, and student performance on standardized tests. These criteria, with corresponding data, were employed in schools near the Rockville, Maryland location of the corporation. This model, following the Rockville success, will be implemented nationwide.

The proposed models offer several unique advantages. First, the models provide a powerful analytical tool for selecting philanthropic projects, where traditional prescriptive models lack a quantitative framework. Second, the models reduce the subjectivity of the selection process to a minimum because a major advantage of the DEA based models presented is that they *do not* require that the relative importance or weights of criteria for selection philanthropic projects be known a priori. Third, in the conventional DEA-based analysis the DMUs are divided into efficient frontiers and others that are not 100% efficient. This study takes an approach opposite to conventional DEA analysis. As we are interested in need assessment, the frontier used in this study is a "need frontier" (not an efficient frontier found in the conventional use of DEA-based analysis) which contains schools with poor performance. Finally, input selections are uniquely assumed as equal to one where none of alternative schools' inputs are under control of donating organizations. The paper both innovates and implements DEA-based models, offering a powerful strategic tool for philanthropic decision-making.

Full references available upon request

REGAL ENTERTAINMENT GROUP: THE MOVIE THEATER INDUSTRY

Robert J. Mockler, St. John's University, 718-990-7419, mocklerr@stjohns.edu
Ho-Sung Lee, St. John's University, 718-990-7415, leeh@@stjohns.edu
Marc Gartenfeld, St. John's University, 480-247-2274, mdj13j@aol.com
Dorothy Dologite, Baruch College, 212-990-6161, Dorothy_Dologite@baruch.cuny.edu

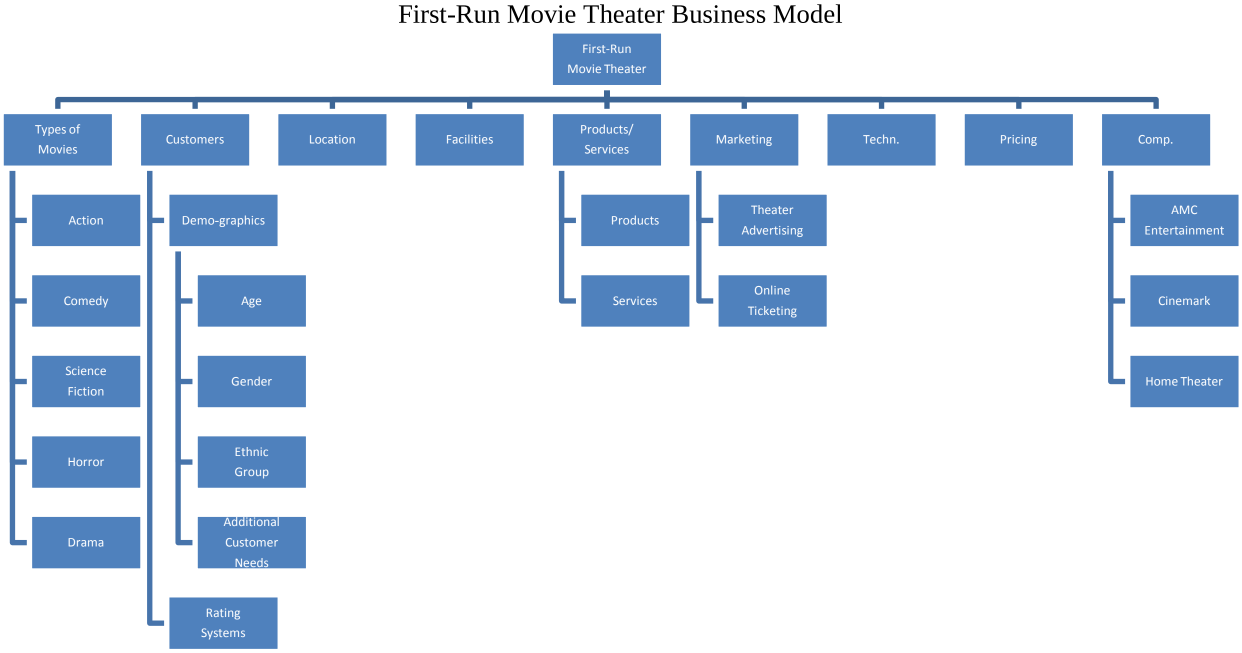
ABSTRACT

Regal Entertainment Group (REG), the US's largest movie theater chain had more than 6,273 screens at about 580 theaters in 40 states through its Regal Cinemas, Edwards' Theatres, United Artists Theatre Company, and Hoyts Cinema brands. During the summer of 2005, Hollywood sank into one of its largest slumps since the 1980s. In December 2005, Michael L. Campbell, the chairman and CEO of Regal Entertainment Group, faced the task of developing an effective differentiating enterprise-wide strategy if REG was to survive and prosper against aggressive competition over the intermediate and long-term future.

INDUSTRY AND COMPETITIVE MARKET: FIRST-RUN THEATERS

The number of movie theater admissions in U.S had been decreasing for the last 2 years (National Association of Theater Owners, 2005). There were several factors that contributed to the decline, including rising ticket prices, customer dissatisfaction with the theater experience, consumers' shift to watching movies sent directly to their mailbox by Netflix or Blockbuster, or to buy movies from cable TV's pay per-view channels, as well as the building of their own movie libraries by purchasing cheap DVDs from mass retailers like Best Buy. Hollywood's obsession with big opening weekend grosses also did not help, since the exhibitor's percentage of a movie's gross increased the longer the film played in the theaters. When a big film was burned out by the second or third week of its release, there was not much business left to fill the theater seats (Hoover, 2005).

Figure 1
First-Run Movie Theater Business Model



Industry Opportunities & Threats and Keys To Success

Tables 1 & 2 summarize the industry's opportunities and threats as well as its keys to success, which, in other words, describe what it takes for a company operating in this industry to be successful in various operational areas.

Table 1
Industry Opportunities And Threats

Opportunities	Threats
Expand new products to gain more market share	Other movie theater can take measures to duplicate strategies similar to REG
Customers require services not offered by other competitors	Limitation of finding new space to expand facilities
New technological advanced provide a greater level of services to customers	Customer trends toward more high quality services
Renovation of facilities provides a great satisfaction to customers	New market trends toward home theaters
Providing advanced online services will offer customers up to date information and convenience	

Table 2
Industry Keys To Success

Area	Keys to Success
Customers	• Have strong moviegoers age groups • Have effective rating systems • Target minority groups • Number of ticketing • Online ticketing • Provide activities such as arcade • Strong Brand Image • Competitive pricing and package plan
Location	• Close to customers • Metropolitan areas • The number of competitors • Population growth • International markets
Facilities	• Proper size of movie theaters • Renovation of facilities • Convenience of facilities and security • Condition of facilities
Products/Services	• Provide Digital Contents • Save cost and storage • Provide various kind of movies • Provide special events • Proper rating systems • Provide up to date movies • Get license for specific movies • Provide various food and drinks • Provide the opportunities of advertisement to merchants • Help in the production of advertisement and contents of merchants • Provide theater rental

Marketing	• Have effective local news paper channels • Convenient telemarketing for customers • Provide the newest movie and theater information • Provide promotion price • Provide up to date its own website • Provide various linked sites to sell its tickets
Technology	• Change analog to digital systems • Hire skilled staff in order to control high technology
Pricing	• Provide different price to different customer group and time • Provide package plan • Provide promotional coupons • Provide membership discount or free services • Provide discount price for group ticketing

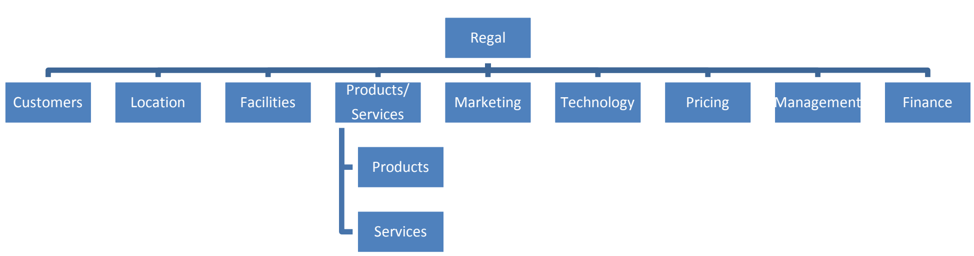
COMPETITION

Regal Entertainment Group (REG) was the US's largest chain theater. The next two largest movie theaters in the USA were AMC Entertainment Inc. and Cinemark, Inc, which were the main competitors of REG. REG also competed against home theaters.

THE COMPANY

Regal Entertainment Group (NYSE: RGC) was the largest motion picture exhibitor in the world. The Company's theatre circuit, comprising Regal Cinemas, United Artists Theatres, and Edwards Theatres operated 6,273 screens in 558 locations in 40 states. Regal operated approximately 18% of all indoor screens in the United States including theatres in 43 of the top 50 U.S. markets and growing suburban markets. The company developed, acquired, and operated multi-screen theatres primarily in mid-sized metropolitan markets and suburban growth areas of larger metropolitan markets throughout the U.S. The company sought to locate each theatre where it would be the sole or leading exhibitor within a particular geographic film-licensing zone. Management believed that as of December 30, 2004, approximately 87% of its screens were located in film licensing zones in which the company was the sole exhibitor (Regal Entertainment Group, 2005). Figure 2 shows the different business segments of Regal Entertainment Group.

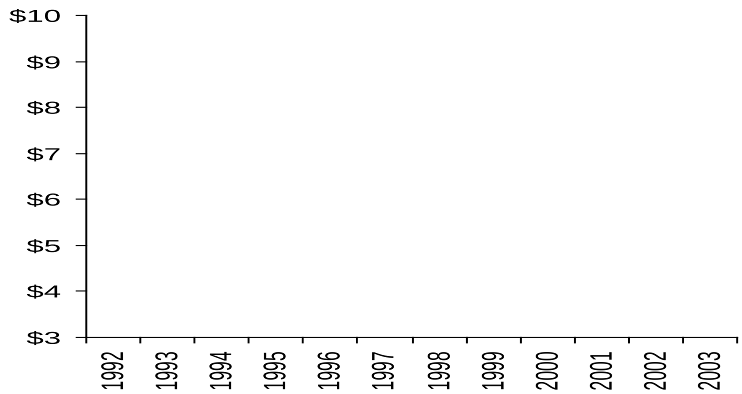
Figure 2
Regal Company



Customers

As shown in Figure 3, although the movie theater industry decreased in their revenue and attendance, REG had increased revenues and attendance (Regal Entertainment Group, 2004).

Figure 3
Steady Box Office Of Regal Growth



Box Office Growth: Avg 6.2% per year

Attendance Growth: Avg. 2.7% per year

Ticket & Concession Prices Growth:
Avg. 2-3% per year

Location

The Company's theater circuit operated 6,273 screens in 558 locations in 40 states. Regal operated approximately 18% of all indoor screens in the United States including theatres in 43 of the top 50 U.S. markets and growing suburban markets. REG developed, acquired, and operated multi-screen theatres primarily in mid-sized metropolitan markets and suburban growth areas of larger metropolitan markets throughout the U.S. As of December 30, 2004, approximately 87% of its screens were located in film licensing zones in which the company was the sole exhibitor. Being the sole exhibitor in a film licensing zone provided the company with access to all films distributed by major distributors and eliminated its need to compete with other exhibitors for films in that zone.

Facilities

REG's theaters housed an average of 11 screens, and more than 60% of its screens were in theaters with stadium seating which defined as a technique used in movie theaters to allow more guests to see the movie screen with less blockage than traditional seating. Like seating in a football or baseball stadium, stadium seating in theaters was usually a 30 degree slope stepped upwards from the bottom of the theater, as opposed to the approximately 15 degree gentle slope in traditional theaters. REG had demonstrated it ability to enhance revenues and realize operating efficiencies through the successful acquisition and integration of 15 theater circuits since 1995. REG had generally achieved immediate cost savings at acquired theater and improved their profitability through the application of its consolidated operating functions and key supplier contracts.

Products/Services (Revenue Model)

As shown Table 3, 67% REG's revenue came from ticket sales and 26% from concessions (Regal Entertainment Group, 2005).

Table 3
Regal Entertainment Group’s 2004 Sales

	$ million	**% of total**

Admissions	1,658	67
Concessions	636	26
Other	174	7
Total	2,468	100

Marketing

REG utilized local newspaper's movie listings, where new releases were advertised with locations of the movie theaters screening the movie at the time. Another commonly used advertising method was utilizing movie phone services which allowed customers to find a movie, show-times, and locations of the movie they wished to watch by calling an automated telephone service. In addition, REG customers could purchase movie tickets online on the Company's own website as well as on third party websites, which were in a joint-venture agreement with REG.

Technology

While REG admitted that digital projection technologies required to screen movies in theaters were not yet commercially viable, the Company was focusing on this lower-cost digital video and communications tool to digitally distribute in-theatre advertising, as well as complement its business of renting out theaters for corporate meetings, seminars, and distance learning programs. Regal had outfitted nearly 80% of its locations with digital projectors and other ancillary digital equipment for marketing and advertising purposes. Regal would not yet digitally project feature-length films, but its system could be fully upgraded once the industry adopted a set of digital standards (Hoover, 2005).

Pricing

The price of movie ticket admissions at REG theaters differed from location to location. For example, in certain cities in California matinee's were priced at $7.50 before 6 P.M. and $9.50 for general adult admissions after 6 P.M. REG also offered senior discount pricing of $7.00. In New York City the Company charged $10.75 for general adult admissions without any discounts for matinees. For senior citizens they offered $7.00 admissions. For children they charged $7.00 in both California and New York City. Joining the Regal Crown Club customers could earn rewards for going to the movies. When customers joined the Regal Crown Club, they received free movies, popcorn, and soft drinks. Regal Crown Club members would receive 1 credit per dollar spent on ticket purchases at the box office, with a maximum of 12 credits per card, per day. All members earned extra credits on concession purchases. Two extra credits were earned for a concession transaction, or 4 extra credits were earned for a specified promotional transaction, such as a Candy Combo. Members received rewards each time 40 credits were earned. REG offered the following pricing discount packages/deals, as listed in Table 4 (Regal Entertainment Group, 2005).

Table 4
Tickets/Packages

Tickets/Packages	Description
PREMIERE SUPER SAVER TICKET	These tickets were accepted for any movie at any time (even opening night of blockbuster films), this unrestricted movie ticket was customers best value at $7.00 each. Plus, customers could add free custom

	messaging on the back of the ticket for a personal touch. A 50 ticket minimum purchase was required
VIP SUPER SAVER TICKET DATE	It was not accepted during the first twelve days of selected new releases, this restricted movie ticket offered the best savings at $6.00 each. 50 Minimum purchase required.
ULTIMATE PREMIERE MOVIE PACK	A pack included two unrestricted Premiere movie tickets and a $10 gift certificate good towards any concession or box office purchase. 25 Minimum purchase required.

Management

REG was the largest domestic motion picture exhibitor with nearly twice as many screens as the nearest competitor. The quality and size of REG theatre circuit was a significant competitive advantage for negotiating attractive national contracts and generating economies of scale. REG had significant experience identifying, completing, and integrating acquisitions of theatre circuits. REG had demonstrated the ability to enhance revenues and realize operating efficiencies through the successful acquisition and integration of 15 theatre circuits since 1995. REG had generally achieved immediate cost savings at acquired theatres and improved their profitability through the application of its consolidated operating functions and key supplier contracts. The company had developed a proven operating philosophy focused on efficient operations and strict cost controls at both the corporate and theater levels. At the corporate level, REG was able to capitalize on its size and operational expertise to achieve economies of scale in purchasing food & beverages and marketing functions. REG had developed an efficient purchasing and distribution supply chain that generated favorable concession margins. At the theater level, management devoted significant attention to cost controls through the use of detailed management reported and performance-based compensation programs to encourage theater managers to control costs effectively and increased concession sales.

Finance

REG had invested over $2.1 billion in capital expenditures since 1997 to expand and upgrade its theater circuit. As a result, REG did not expect to require major capital reinvestments in the near term to maintain its operations. The combination of its operating margins and its limited need to make maintenance capital expenditures would allow the company to generate significant cash flow from operations. For the thirty-nine week ended 29 September 2005, Regal Entertainment Group's revenues increased 1% to $1.85B. Net income decreased 2% to $56.7M. Revenues reflected vendor marketing programs and the purchase of theaters. Net income was offset by higher film rental & advertising costs, an increase in interest expense and the presence of equity in earnings joint venture including former employee compensation expenses (Regal Entertainment Group, 2005).

LOOKING TOWARDS THE FUTURE

Michael L. Campbell, the chairman and CEO of Regal Entertainment Group, was faced with many crucial decisions affecting the future of his theater. One very important decision involved the future locations of the Company's theaters. The location decision revolved around two alternatives: to expand to foreign regions, such as Asia, or to existing domestic high population and high-income areas.

References upon request

An Experiential Learning Case: A University-Based Field Trip to the Racetrack

Jack Rappaport
Management Department
LaSalle University
Philadelphia , PA 19141
E-mail: rappapor@lasalle.edu
215-991-3555

Stephen B. Richter
Department of Accountancy and MIS
Villanova University
Villanova, PA 19085
E-mail: stephen.richter@villanova.edu
215-643-0163

Abstract

This paper extends the research of Rappaport and Richter [28] in which they describe an experiential approach to decision making and statistics using the racetrack betting markets. In this paper we extend the implementation of the experiment to include a full blown field trip to the actual racetrack. We describe the results of the field trip that was implemented in the spring of 2009 and explain how it impacted the educational process beyond the classroom exercise.

I. Introduction

This paper extends the research of Rappaport and Richter [28] in which they describe an experiential approach to teaching decision making and statistics using the racetrack betting markets. In this paper they show how the class can be transformed into a mini-simulcast center, whereby the races are simulcast from the Internet and projected onto a computer screen. The experiment can take place in a computer lab whereby each student would have access to his or her own computer so that he or she can monitor the betting patterns of the races of his or her own choice and use a series of Excel spreadsheets to monitor the results by generating the appropriate statistical results. This experiment can be beneficial in either a statistics or a decision making class.

In this paper we extend the implementation of the experiment to include a full blown field trip to the actual racetrack. We describe the results of the field trip that was implemented in the spring of 2009 and explain how it impacted the educational process beyond the classroom exercise. The paper also considers how some of the initial resistance from some of the university officials were overcome within the context of some of the theories of change management. Finally, the paper proposes how future field trips can be structured to further enhance the educational attributes and can be used to develop a possible funding source from the horse racing industry.

1. Experiential Learning

Just as in the case of the classroom experiment, the purpose of the field trip is to implement the concept of experiential learning that has been widely accepted as a useful approach to learning in many academic fields. Experiential learning is learning through reflection on doing, which is often contrasted with rote or didactic learning. An example of experiential learning is going to the zoo and learning through observation and interaction with the zoo environment as opposed to reading about animals from a book. Thus one makes the discoveries and experiments with first hand knowledge, instead of reading from a book. Experiential learning theory draws on the works of prominent 20th scholars who gave experience a central role in their theories of human learning and development; notably John Dewey, Kurt Lewin, Jean Piaget, William James, Carl Jung, and others. For example, John Dewey [18] conceived of education as a continuing reconstruction of experience, whereby the process and goal of education are one and the same.

The use of experiential learning techniques has been shown to be important in the teaching of statistical concepts. Active learning experiences can provide students with valuable oppportunities to apply theory and practice ([2], [3], [13], [15], [32], [43]).

This paper extends the experiential approach to both statistics and decision making classes to engage the students in a real world experience by taking them to the racetrack. Many researchers have highlighted the importance of social interactions or the "human moment" to educational outcomes. Experiential learning techniques may allow students to experience the more intangible aspects of institutions such as a sense of inclusion or community. The concept of the racetrack field trip can be seen within this philosophy of education, since it allows the students to experience the concepts of risk and return within an extended social setting; the concepts of risk and speculation are made more apparent as the students are immersed in a real gambling environment.

2. Academic Research versus Field Research

The field trip to the racetrack makes explicit the dichotomy between the academic approach and the real world approach. This duality can best be understood by comparing the academic research with the more popular literature that is read by the general public.

The academic research can be separated into two major groups; the first group deals primarily with the development of decision strategies for making horses race selections.

Many of the models are very sophisticated from an academic standpoint since there has been a significant contribution to this field from many high quality academic researchers ([8], [16], [33], [31], [40], [26], [11], [19], [29], [37], [41]).

Another group of articles are concerned with the study of risk; virtually all of the empirical work analyzing risk preferences in the horse race betting markets shows the representative bettor to be risk loving. This has resulted in what is called the "favorite-longshot" bias; an established feature of the betting markets whereby longshots win less often than their subjective probabilities imply and favorites win more often; in other words bettors tend to prefer long shots to favorites because they are motivated by the possibility of a large payout. There has been a considerable amount of research studying this phenomenon ([41], [35], [1], [27], [34], [39], [42], [30], [36], [12]).

On the popular side, there are a large number of publications and information sources available to the average horseplayer to help them better understand the sport and make better decisions. These kinds of publications would be considered trade publication by most academics. The most important source of information is the Daily Racing Form (DRF), a newspaper-style publication which can be an important tool for the handicapper or horseplayer; the DRF provides a large amount of statistical information about each horse including detailed past performance lines, lifetime records, amount of money earned, and a myriad of other pieces of information that can be used for either casual or serious study. There are also numerous books on the subject that provide a variety of theories on how to best interpret the information provided by the DRF. For example a number of well known handicappers such as Andrew Beyer ([5], [6], [7]) have written books on this topic. The DRF also publishes books that study special aspects of horse race handicapping, such as how to interpret such factors as speed and pedigree ([20], [38], [21], [9], [18], [17], [10], [25]).

One can contrast and compare the academic works with the trade publications to better understand the dualities that exist between the academic world and the real world; one can consider them as studying the same phenomena using a different perspective or approach. The one is more theoretical and scientifically rigorous whereas the other is more anecdotal but more real world oriented. Most academic researchers would probably view the popular literature as unscientific and superficial, whereas the average horseplayer would probably view them as unpractical, even if they could understand them. From an holistic point of view one might consider each of these approaches as having the same goal but approaching the problem from a different perspective. The dissemination and processing of knowledge can be seen in relative terms and dependent on the point of view of the audience that the knowledge is intended for; this dichotomy can be seen as well in the experiential format in which we plan the racetrack field trip. The academic side can be enhanced by the practical side, and vice versa; each approach can be considered a different way to achieve the same goal.

3. The Initial Field Trip

The elements of the field trip are essentially the same as the ones for the classroom experiment as explained by Rappaport and Richter [28], but they would be adapted to the racetrack environment. First we review these basic aspects as presented by Rappaport and Richter [28]. One of the basic aspects of the experiment is to have the students use and develop a series of Excel spreadsheets that generate a variety of statistical reports relating to the betting process. The following spreadsheet generates the mean return per dollar bet and the standard deviation per dollar bet for a typical betting session; other spreadsheets can be developed to generate a variety of other statistical results such as the one that shows the 99% confidence limits for the mean return per dollar bet for different sample sizes (i.e. the number of horses wagered); note that the spreadsheet programs could be (at least in principle) implemented at the racetrack with the use of laptop computers.

FINAL ODDS	AMOUNT WAGERED	WIN/LOSS (O or 1)	TOTAL RETURN	RETURN (per $bet)	SUM OF SQUARES
2.50	$10.00	1.00	$25.00	2.50	34.60
5.00	$1.00	1.00	$5.00	5.00	19.01
1.20	$10.00	0.00	-$10.00	-1.00	3.14
3.50	$1.00	0.00	-$1.00	-1.00	8.18
2.80	$1.00	0.00	-$1.00	-1.00	4.67
3.20	$1.00	0.00	-$1.00	-1.00	6.55
5.50	$1.00	0.00	-$1.00	-1.00	23.62
	$25.00		$16.00		99.76

TOTAL WAGERED	TOTAL PROFIT	MEAN RETURN (per $bet)	STANDARD DEVIATION (per $bet)
$25.00	$16.00	$0.64	$2.04

These results can be represented by the following formulas:

O_i = odds of wager i (n= the total number of wagers)
V_i = 1 if wager i wins
V_i = 0 if wager i loses
W_i = the amount of money bet on wager i
E = mean return per dollar bet
S = standard deviation per dollar bet

$$E = \frac{\sum_{i=1}^{n} \left(W_i V_i (O_i + 1) - W_i \right)}{\sum_{i=1}^{n} W_i}$$

$$S = \sqrt{\frac{\sum_{i=1}^{n} W_i \left(V_i (O_i + 1) - 1 - E \right)^2}{\sum_{i=1}^{n} W_i - 1}}$$

Another aspect of the experiment concerns the type of strategy used by the decision makers. Two overall approaches are considered- rule based versus intuitive or judgmentally based decision strategies. A rule based system would give the same results independent of the individual preferences of the decision makers since the decisions would depend on a predetermined set of rules that would specify what horse or horses to wager on. The rules could be based upon the information provided in the past performance lines (such as choose the horse with the highest speed rating in the last race) or the betting patterns (such as choose the horse whose odds have the greatest percentage difference below the morning line odds). An intuitive system would rely primarily on the way the decision makers interpret the data and would ultimately depend upon their own preferences (as well potential biases). Intuitively based systems would most likely work the best when the decision makers have a great deal of experience handicapping the races and understand some of the finer subtleties of the sport (although sometimes good decisions can be made by newcomers to the game, i.e. the notions of beginner's luck).

Finally, the experiment considers the propensity of the decision maker for risk. Those with a low propensity for risk would tend to play the favorites or lower odds horses and those with a higher propensity for risk would tend to play longshots or horses with relatively high odds. The final choices of the decision makers would result from an interaction of their risk preferences and their decision making strategy.

The field trip (at least initially) would be less structured than the classroom experiment. Students would be allowed to wander around on their own, form small groups on an ad hoc basis and basically try to have a casual day at the track. They would be free to explore the racetrack environment and gamble as they choose, but would be asked to eventually report back (either individually or as a group) on the results of their experience based upon the principles outlined in the classroom case.

Although the elements of the field trip would be primarily based upon the classroom case it would have many aspects that would make it significantly different from the classroom case that could potentially add some new dimensions to the learning experience. These are listed below:

a. The attendance of the students at the racetrack would increase the likelihood that they would make real, as opposed to mock wager. The process of making real wagers would change the character of the experiment as it could potentially transform the experiential learning case into an example of action learning, whereby the participants study the results of their own actions and experience. In addition, the students that were over 21 would be permitted to gamble in the casino which was at the time in the same building as the racetrack. The experience of casino gambling would extend the notion of risk and uncertainty in other ways.

b. The decision process at the racetrack can be different than the decision making process of on-line wagering for a variety of reasons. Observing the actual horses in the paddock or post parade before the race gives the decision makers more information to utilize. Also it is much easier to change one's bet at the track than wagering on-line. The potential interaction with other gamblers gives another dimension to the wagering process at the track as opposed to on-line. The live environment also gives one a much better sense of the tensions that build up before the race and thus makes more explicit the conflicting forces of information and knowledge that lead to the decisions. The ability to watch the race live and in person can add another dimension to his or her experience and therefore can potentially affect the decision making process itself.

c. The field trip can help the students better understand the difference and similarities between the academic world and the real world side as it pertains to the study of horse racing, or for that matter any business process or sport. Watching real horse players make decisions complements the study of the same process from an academic point of view. Thus the concepts of risk and decision making can be looked at from both points of view.

5. Gaining the Acceptance of the Stakeholders- An Example of Change Management

In this section we report on the process by which the acceptance of the trip was obtained from the various stakeholders and university officials. There was an expectation that there would be some initial resistance to having the university officially sponsor the trip. It should be mentioned that the organizer of the trip (who is the first author of this paper) has a reputation at the university of creating somewhat controversial programs and being very entreprenurial in spirit but he has been being fairly successful over the years of overcoming resistance from the university to many of his ideas. For example, many years ago he was successful in teaching a course in casino management, whereby he applied the concepts of operations management it the casino industry as well as inviting a number of fairly high level casino executives as guest speakers. He carried out this plan along with a major placement of students to work in the Atlantic City casinos during the summer after overcoming a fair amount of resistance from a variety of university officials. At this point the idea of creating some form of a casino management program has become much more acceptable, although a full blown program would require a formal initiative on the part of the university administration.

Now we list some of the stakeholders and their initial perspectives on the idea of the field trip:

a. The Dean of the Business School

The Dean of Business was generally positive about the trip, although he seemed initially cautious, expressing his concerns about the idea of taking a group of undergraduate students to a gambling establishment where they would also be allowed to drink (assuming they were 21). Although he seemed willing to approve the trip on his own, he seemed interested in finding out the opinions of some of the other university officials such as the Dean of Students. The academic relevance of trip did not seem to be much of an issue for him.

b. The Business School Faculty

The organizer of the trip is a tenured faculty member with over 20 years of service to the university who has the reputation of being somewhat of a “loose cannon”. He generally seems to have gained the respect and friendship of many of his colleagues over the years. Although the faculty did not play any formal role in obtain the approval for the trip, their opinion clearly would have an informal influence on the Dean. The track record of the organizer in terms of research and teaching in this subject clearly created a positive outlook on the part of the faculty to the idea of the trip.

c. The Dean of Students

The Dean of Students was much less receptive to the idea of the trip, although she deferred the responsibility to the Dean of the business school. It seemed apparent that although she wasn’t going to do anything to stop the trip, she wasn’t going to provide any particular support or encouragement either. Although she stated that she really didn’t have to be consulted on this matter at all (leaving the total responsibility of the trip to the Dean of the business school) she nevertheless insisted on a fairly beauracratic and onerous process of signing the students on and off the bus, refused to allow students from other schools on the bus, and in general took a rather negative view on the concept of the trip. Finally, she insisted that the organizer consult with the NCAA compliance officer for the possibility that atheletes would not be allowed to gamble because of a possible violation of the NCAA rules on gambling.

d. The NCAA Compliance Officer

The NCAA compliance officer was probably the most resistant to the trip since she was very concerned on how it might appear to the officials of the NCAA which has strict rules on the involvement of gambling by student atheletes. The rules allow the student athletes to gamble on any game or sport that is not sponsored by the NCAA, as long as they are of the legal age, so that, for example, they would not be allowed to gamble on a football or basketball game, but they could gamble on poker or horse racing (it should be noted that the legal age for horse race gambling is 18). However, there was an informal tone to the rules that strongly discourages any form of gambling because of its addictive nature and the possibility of it leading to the types of gambling that would be strictly prohibited. For this reason, including perhaps her own personal biases on the matter, the compliance officer tried to create some roadblocks for the trip at least as far as the atheletes were concerned. She wrote a fairly formal letter to all of the atheletes that would potentially go on the trip that they would officially be allowed to participate but showed strong reservations about their involvement in gambling. At the very least, the compliance officer displayed a very negative reaction to the idea of the trip.

e. The Students

The students were probably the least resistant to the trip, which was probably not surprising since many of them probably thought of the trip as a way to have an enjoyable afternoon and get away from the school. The organizer made it even more appealing to them by making the trip purely voluntary, with the promise of a significant amount of extra credit points should they participate. Those that not did not choose to participate would be eligible to earn the same amount of extra credit points by either carry out a short paper or case on the races or a similar topic. Needless to say, a significant number of students signed up for the trip, although the attitude of the compliance officer did initially telegraph a negative image to the atheletes which may have also been transmitted to the general student.

Although the organizer of the trip didn’t see any major stumbling block to the project, he wanted to smooth out any potential conflicts or reservations and create a situation whereby he obtained a full level of support from all the interested parties- both formally and informally. He achieved this objective by trying to avoid confrontation as much as possible and communicated with each of the stakeholders to gain their ultimate support and respect. The key to this process was his willingness to send a list of all of the athletes that would potentially participate on the trip to the compliance officer, who then responded by sending them a fairly formal letter. The goal was to “neutralize” the position of the compliance officer in order to gain a more solid base of support from the Dean and the faculty which in turn increased the potential interest on the part of the students. Actually one of the catalysts for the plan happened almost accidentally, whereby one of the supervisors of accounts payable, who was a former student of the organizer of the trip suggested that the purchasing department could rent a bus at a vastly reduced price and this enabled the organizer to obtain the entire funding of the bus from the racetrack. This had a very positive impact on the Dean, who as most Deans are, was very cost conscious. Once it was clear that the bus was officially ordered through the university, most of the other elements of resistance to the trip quickly faded.

The process by which the change in the overall organizational attitude was accomplished can be appreciated by one of the comments from a relatively low-key university administrator who said was quite impressed by the efforts of the organizer to gain the official support of the university to sponsor a bus trip to the casino-racetrack as part of the academic experience.

The process described above can be viewed in light of the many publications on change management that give practical approaches to managing and implementing change. John Kotter's highly regarded books, "Leading Change" [23] and the follow-up "The Heart of Change" [24] describe a model for understanding and managing change. Each stage acknowledges a key principle identified by Kotter relating to people's response and approach to change, in which people see, feel, and then change. A more detailed interpretation of the personal change process can be seen in John Fischer's model of the process of personal change.

Kotter developed an 8 step change model for the effective of change (increase urgency, build the guiding team, get the vision right, communicate for buy-in, empower action, create short term wins, don't let up and make change stick). Steps 4 and 5 (communicate for buy-in and empower action) seemed to be key in the process described above as the organizer empowered the action of the students by rewarding them for participating, allowed the compliance officer to communicate her goals to the students, and took advantage of the participation of the purchasing department to order the bus.

6. The Results of the Field Trip

The results of the experiment showed that the students were more interested in the overall notion of risk and return than in the decision strategies. Thus the students were not so much interested in horse racing as an end in itself, but rather its relevance to the concepts of risk and reward. The main focus of the students seemed to be the issue of whether to play favorites or longshots, monitoring the mean return per dollar wagered, and comparing their performance with other student. The concept of statistical sampling and confidence intervals also played a role in their thinking. In fact, the ability of the trip to motivate the students to think in terms of risk and reward has prompted many of the finance professors to view the trip as very educational within the context of speculative markets in general.

We describe on anecdotal case to illustrate how the experiment has the capability of engaging students in a process of scientific inquiry, using the notions of risk and decision making. One of the students made a fairly large show bet on a 25-1 horse that came in second. He consequently made a considerable profit on this wager and actually won money over the entire using this strategy and naturally became interested in studying the result using this type of strategy over a long period of time. Knowing that there was a favorite-longshot bias for win bets, it was natural for him to consider the possibility of betting on the longshots for show. Although he had made a profit for the day, he quickly realized that his strategy would probably not work in the long run, primarily because it would violate the efficient markets hypothesis. Nonetheless, he seemed interested to investigate this issue further and perhaps modify the strategy with a particular selection process based upon the past performance line. Perhaps his theory may seem a little naïve to a seasoned racetrack player, but it does show how a student can be motivated to think in scientific terms and at least in principle understand the value of carrying out the scientific method.

7. The Field Trip as an Experiential Case

Kolb and Kolb [22] outlined six propositions to be the fundamental basis for an experiential learning process (that learning is conceived of as a process, not in terms of outcomes; that all learning is relearning; that learning requires the resolution of conflicts between dialectically opposed modes of adaptation to the world; that learning is a holistic process of adaptation to the world; that learning results from synergetic transactions between the person and the environment; and that learning is the process of creating knowledge). Rappaport and Richter [28] showed how the classroom experiment seemed to successfully illustrate each of these propositions. We believe that the field trip, as an extension of the classroom experiment, illustrates these propositions as well, but in a more action oriented environment. In particular, the information that the students learned in the classroom cannot be totally divorced from what they experienced at the track. The understanding of the concepts of risk and return cannot be separated from the excitement gained from the live race, or any other social or emotional forces that they might experience at the track. In a similar manner, the study of the stock market cannot be isolated from the real financial environment which has its own set of social and emotional forces.

Finally, we claim that the experiment can give students an insight into the notion proposed by Zeleny [44] that decision making refers to the process of selecting an option or alternative and taking full responsibility for the consequences. The field trip is a good way to illustrate this principle since it implicitly encourages the students to make real wagers and feel accountable for the results.

8. The Structure of Future Field Trips

Future field trips are currently being planned and more structured approaches will be considered. One of main idea is to develop a handicapping contest between groups of students and perhaps between students of different schools. Plans are currently underway to implement such a plan which would add a more structured element and further enhance its educational value.

The idea of developing a contest for students is not new, and has in fact has become a very prestigious kind of event for business schools to engage their students in. These competitions, usually called case competitions, are generally sponsored by an outside organization such as a corporation or an accounting firm and require the students to develop and present the results of a business case to a panel of judges. The winners of the competition at a local school go on to compete against the winners from other schools and some prize money is usually awarded based upon the results of the competition.

This concept can potentially be extended to the racetrack application. Racetracks generally sponsor handicapping contests for their customers on an ongoing basis. This concept would be simply be adapted to university students and could be tailored to either undergraduates, graduate students, or a combination of both.

This plan would also open up the possibility of obtaining outside funding for the general racetrack initiative. Not only the racetrack, but also any horse owner or breeder could be a possible funding source for this kind of program. Horse racing has been confronted by rather stiff competition for the gambling dollar from competing forms of gambling such as casinos and state lotteries. Many in the industry are concerned about the long term viability of the sport particularly since it has historically drawn its customers from the ranks of the older population. There are many very wealthy individuals who have effectively become patrons of the sport because of their involvement either as horse owners, horse breeders, or both. Many of these individuals could become potential funding sources for any educational projects relating to horse racing.

References

[1] Asch, P., Malkiel, B.J., and Quandt, R.E., (1984), "Market Efficiency in Racetrack Betting", The Journal of Business, Vol. 57, No.2, April

[2] Akram,M., Siddiqui,A. J., and Yaseem,F., (2004) "Learning Statistical Concepts", International Journal of Mathematical Education in Science and Technology, Vol 35 No 1

[3] Anderson,E., "Active learning in the lecture hall" (1997) Journal of College Science Teaching, Vol 26 May

[4] Bain,R.S.,Hausch,D.B. and Ziemba (2006) "An application of expert information to win betting on the Kentucky Derby, 1981-2005", The European Journal of Finance, Jun, Iss.4

[5] Beyer,A. (2007) The Winning Horseplayer:An Advanced Approach to Thoroughbred Handicapping, Houghton Mifflin

[6] Beyer,A. (1994) Picking Winners:A Horseplayer's Guide, Houghton Mifflin

[7] Beyer,A. (1995) Beyer on Speed:New Strategies for Racetrack Betting, Houghton Mifflin

[8] Bolton,R.N. and Chapman,R.G. (1968) "Searching for Positive Returns at the Track: A Multinomial Logit Model for Handicapping Horse Races", Management Science, Aug, Vol 32 Iss 8, Linthicum

[9] Borg, N., Off The Charts (2003) Daily Racing Form Press

[10] Cardello, J., Speed to Spare (2003) Daily Racing Form Press

[11] Chezum,B. and Wimmer,B.S. (2000) "Evidence of adverse selection from thoroughbred wagering", Southern Economic Journal, Stillwater, Jan, Vol66 Iss 3

[12] Coleman, C. (2004) "New light on the longshot bias", Applied Economics" London, Mar Vol 36 Iss 4

[13] Dabbour,K. (2001) "Applying active learning methods to the design of library instructions for a freshman seminar", College and Research Libraries, vol 58 May

[14] Dewey,J. (1938) Education and Experience, New York: Simon and Schuster

[15] Hakeem,S. (2001) “Effect of Experiential learning in Business Statistics”, Journal of Education for Business, Nov-Dec

[16] Hausch,D.B. and Ziemba,W.T. (1990) “Arbitrage Strategies for Cross-Track Betting on Major Horse Races”, The Journal of Business, Chicago Jan, Vol 63 Iss 1

[17] Heller, B. (2004) Go For The Green, Daily Racing Form Press

[18] Illman, D., Betting Maidens & 2-Year-Olds (2005) Daily Racing Form Press

[19] Iyengar,G.N. and Cover,T.M. (2000) “Growth optimal investment in horse race markets with costs”, IEEE Transactions on Information Theory Nov, New York

[20] Keppler,D.,Trainer Angles (2006) Daily Racing Form Press

[21] Klein, S., The Power of Early Speed (2005) Daily Racing Form Press

[22] Kolb, A.Y. and Kolb, D. (2005) Learning Styles and Learning Spaces: Enhancing Experiential Learning in Higher Education, Academy of Management Learning and Education, Vol 4, No. 2

[23] Kotter, J. (1996) Leading Change, Harvard Business School Press

[24] Kotter, J. (2002) The Heart of change, Harvard Business School Press

[25] Michaels, N. (2005) Handicapping Contest Handbook Daily Racing Form Press

[26] Phatarfod,R. (1999) “Betting strategies in horse races”, Asia-Pacific Journal of Operational Research, Singapore May, Vol 16 Iss 1

[27] Quandt, R.E. (1986) “Betting and Equilibrium”, The Quarterly Journal of Economics, Vol. 101, No. 1, Feb

[28] Rappaport, J. and Richter, S.B. (2009) “Teaching Note: Decision making using the racetrack betting markets”, International Journal of Applied Decision Sciences

[29] Rosenbloom,E.S. (2003) “A better probability model for racetrack using Beyer speed numbers”, Omega Oct, Vol 31 Iss. 5

[30] Rhoda,K.L.,Olsen, G.L., Rappaport,J.M. (1999) “Risk Preferences and Information Flows in Racetrack Betting Markets” The Journal of Financial Research, Vol XXII, No 3, Fall

[31] Schnytzer,A. and Shilony,Y. (1995) “Inside information in a betting market”, The Economic Journal, London, Jul, Vol 105 Iss 431

[32] Sharkarian,D. (1995) "Beyond lecture: Active learning strategies that work", Journal of Physical Education, Recreation and Dance, Vol 66, May-June

[33] Shin,H.S. (1991) "Optimal Betting Odds Against Inside Traders", The Economic Journal, London, Vol 101 Iss 408

[34] Shin,H.S. (1992) "Prices of State Contingent Claims with Insider Traders, and the Favorite-Longshot Bias", The Economic Journal, London, Mar, Vol 102, Iss 411

[35] Snyder,W. (1978) "Horse Racing: Testing the Efficient Market Hypothesis", Journal of Finance, 33

[36] Sobel,R.S. and Raines,S.T. (2003) "An examination of the empirical derivatives of the favorite-longshot bias in racetrack betting", Applied Economics, London, Mar 4 Vol 35 Iss 4

[37] Stein,W.E. and Mizzi,P.J. (2003) "An analysis of exotic wagers in a parimutuel setting", Applied Economics, London, Mar 4 Vol 35 Iss 4

[38] Stitch, L. (2004) Pedigree Handicapping Daily Racing Press

[39] Swindler,S. and Shaw,R. (1995) "Racetrack wagering and the "uninformed" bettor: A study of market efficiency", Quarterly Review of Economics and Finance, Greenwich Fall 1995 Vol 35 Iss 3

[40] Thalheimer,R. and Ali,M.M. (1995) "Exotic betting opportunites, pricing policies and the demand for parimutuel horse race wagering", Applied Economics, London Aug, Vol 27 Iss 8

[41] Weitzman,M. (1965) "Utility Analysis and Group Behavior: An Empirical Study", Journal of Political Economy, 73

[42] Williams, L.V. and Paton,D. (1997) "Why is there a favorite-longshot bias in British racetrack betting markets?", The Economic Journal, London, Jan Vol 107 Iss 440

[43] Zeis,C., Shah, A., Regassa, H., Ahmadian, A. (2001) "Statistical Components of an Undergraduate Business Degree: Putting the Horse before the Cart", Journal of Education for Business, Vol77, No2, Nov-Dec

[44] Zeleny, M. (2008) "The KM-MCDM interface in decision design:tradeoffs-free conflict dissolution", International Journal of Applied Decision Sciences, Vol 1, No 1

TIRE MANAGEMENT OF GE TRAILER FLEET SERVICES

O'Brien Sean, Richard Stockton College of New Jersey, spob15@aol.com
Shaoping Zhao, Richard Stockton College of New Jersey, szhao@stockton.edu

ABSTRACT

This case concerns General Electric's Trailer Fleet Services business's needs to manage its largest expense category, tires. The first portion will deal with a brief overview of what GE Trailer Fleet Services is and what the shortcomings of the current process are. The remainder will focus on how GE intends to solve the problems using management information systems, and process improvement.

GE TIRE MANAGEMENT

Trailer Fleet Services (TFS) was founded in 1957 as a transportation equipment rental and leasing company named Transport Pool. It has since become a $550 million revenue unit in the Equipment Services division of the General Electric Company. The small business has grown to become the largest trailer leasing business in North America now one of the most diverse fleets with over 100 branches and more than 135,000 (eight tires/trailer) trailers on the road. The business has expanded from the core rental and leasing in to a "total fleet management" proposition. GE Trailer Fleet Services offers products for every part of the equipment life cycle. From financing new equipment, maintenance of rental or customer owned equipment, to satellite tracking and resale opportunities GE Trailer Fleet Services has grown to meet the needs of a diverse industry.

The lifecycle of a tire is long. A properly maintained tire has an almost infinite life. When I speak of tires there are two main aspects, new or original and retreaded tires. A new tire is just as it sounds, a tire that has never been used has two parts, the tread, and the casing. A retreaded tire is the casing of an original that through a manufacturing process has a new rubber tread installed. Retread manufacturers use various equipment including MRI machines to examine casings for defects, so as long as a tire passes the process it should be serviceable. TFS follows industry recommendations, which say a casing should be destroyed when it is over seven years old or has been retreaded three times.

The largest change for TFS will be treating the casings as an asset. Since the casing is reusable, it has a value that changes with the tires age. Current processes state that the branch maintains its own "pool" of tires. This means the branch is responsible for tires on domiciled equipment and replacement tires. There is no centralized inventory and the entire process is manual. Branches use a printed Excel spreadsheet and a binder to manage their inventory.

Because the business is decentralized by nature, most branches only worried about achieving local optima. A recent change in business structure and incentive plans helps to move away from the "my own business" model. Branch management is now part of the larger business, and has a better use and understanding of the entire system.

Purchasing is conducted in two main ways the Home Run system; which is a dated ERP system with no "part" inventory capabilities, and EM-Smart, a maintenance management system.

Both systems are essentially PO issuing systems with no line item detail. The only visibility to the purchase outside the branch is PO amount, and vendor name, date, etc. There is no specific information on tire size, tread design, or other important features. When sourcing product corporate only knows how much was spent with a particular supplier, even exact quantity information is directional at best.

The limited visibility to spend is the first major shortfall in the current process. It has impacted negotiations, product uniformity, and service levels. One of TFS's largest detractors is inconsistency at the branch level. A customer may have tire work done in the same period at several different branches and receive several different tires, and prices. Efforts have since been made to improve consistency and standardize to certain products, with the help of vendors, and purchase controls. While this has improved things slightly we still rely on our vendors to tell us what we purchased.

The other major shortfall in the current process is lack of inventory management. Currently the branches manage locally and use a paper system. There is no central system with branch tire inventories. This lack of visibility impacts several areas of the business. Finance has trouble getting accurate inventory estimates, maintenance leadership does not get an accurate view of cost per mile, and business leaders do not get a clear picture to make better decisions. It is also impossible to search for trends, and evaluate performance.

Two years ago TFS implemented a parts purchasing and inventory system they call Parts Inventory (PI). This system is a web-based application that sends orders to suppliers, and manages inventory. The inventory piece has never been implemented, and tires were not initially included in the ordering process because of "difficulty" in ordering.

When ordering a retread tire, they are like snowflakes. During the retread process analysis finds things needing repair. One tire might need a nail hole patched, plus the tread; another may need a different sort of patch. All the ancillary costs are different, and system requirements require a match between the PO issued and the invoice received. If the amounts don't match, the system rejects it, which delays payment to the supplier. In short the order entry person must know exactly what costs are associated with the tire.

The inventory potion has not been used because of implementation expense. Implementation was piloted in several branches of varying size. The process involved taking a complete physical inventory and labeling and bar-coding locations and parts. The branch was not to be shut down while this occurred. Corporate staff traveled in to conduct the inventory and manage the implementation. In the smallest branch it took eight people 5 full days to get the branch up and running, this is with minimal branch training. The larger branch took two weeks. Implementation was "back-burnered" for cost reasons.

The business's operational leadership decided to evaluate using the full PI system for tires only. Unlike many vehicle parts, tires are easily identifiable and countable. Tires make up the largest expense replacement part for TFS an estimated $10 million annually. TFS had to do something to manage this much spend.

Field management needed visibility into their branch spend. The paper system was not conducive to quick decisions. If a branch needed a history they had to comb through several files. The process was also very time consuming. The new information would allow everyone to know exactly how profitable a customer is.

The team put together to perform the project comprised of managers from several departments including: sourcing, IT, maintenance, and operations/fulfillment. The first of several lengthy workout sessions established the need for user experts, so several branch staff employees were included in the project. They have first hand knowledge of the process and how it can be effectively implemented.

The project started by gathering real world examples of purchases and receipts. The team even used model tires to simulate tires moving through the process. Adam Gianetti, the business maintenance and analytics leader said: “We essentially locked ourselves in a conference room for six hours at a time. It is the only way to get away from the phone and email to get the tasks done”. “Using the model tires seemed a little silly at first, but they were a great tool to follow the process. If something dropped out, we could see exactly where it happened”.

The bulk of the IT work was to enter all the part information into the PI system, and build the databases. This information is the line item detail for each vendor, sizes, tread patterns, and ancillary repairs, etc. The pricing and “part” information was entered into the contacts part of the system since all tires “should” be purchased through a national account provider.

Determinations also had to be made about what should be inventoried and what should not be. When assigning a tire to a work order, the tire is added not the tire plus all the added items. Instead of having a mechanic add eight separate lines for one tire, it was decided that only the casing would be inventoried, as casings are the part of a tire that are an asset and have a renewable life. So orders would be processed and the system would use the average cost associated with a tire. There are two inherent dilemmas with this approach. Order entry is a time consuming task, the team must make efforts to simplify this process. Also, if only the casings are inventoried, that means that the cost assigned to a work order will be the average cost of a casing, which is only about 40% of the total tire.

TFS uses a mounted wheel program; this means that all tires are mounted rims at the vendor location. This is mostly due to safety concerns. So, when a tire is ordered the entry person must enter a tread, casing, repairs, valve stem, and mounting. For new tires we created tire “kits” in PI. There are no variable repairs on a new tire; only the tire, mounting, and valve stem. This simplified new tire ordering to one order entry. In order to simplify the numerous part numbers to remember the team created a quick access card that they could keep at the computer. The card will also include cost information so the branch can accurately do job costing.

Ordering in PI allows central access to any information about the purchase. Sourcing now has the ability to forecast spend, and scrutinize purchasing down to the user level. The inventory management piece has the biggest impact on the business. Once the order is received in PI it is added to the inventory. The inventory has RF and barcode reading capabilities. The road mechanics use a handheld device, which uses a cellular signal to keep inventory updated in real time. As an added benefit the manager can cross reference the trucks GPS system to see if a certain tire is nearby thus avoiding a trip to the shop or vendor location.

With the inventory linked to the maintenance system the business has visibility to all tire usage. Using information gathered during inbound and outbound inspections, and during in-service repairs cost per mile could be analyzed. This provides insight into customer usage, and overall costs. The customer can be compared to others in the same segment, and ensure they are maintaining and spec’ing their equipment properly.

One customer was going through tires and an accelerated rate. The new information allowed the local staff to change the tread selection. The standard tire for a long haul trailer did not work for this customer who did more city driving. The branch was able to monitor the tires and change the tread pattern choice. This instantly lengthened the tire life, saving TFS on replacement costs and administration. The customer reduced equipment downtime, and improved fuel economy.

After the team walked through the process several times all the while taking notes, a process map and policy manual were completed. The next steps was piloting in select branches. Piloting was to be conducted in five branches of varying size and market. There is also some difference in how branches generate revenue. Some branches have great rental and lease revenues, while their revenue from maintenance services is low; other branches are the opposite, they don't rent much, but have 28 mechanics on the road everyday.

Piloting was to be conducted in the following locations:

- Springfield, MA – Large mixed use fleet, and strong maintenance operations
- South Kearny, NJ – Strong rental and lease but limited maintenance
- San Diego, CA – Small branch with heavy border traffic
- Detroit, MI – Largest maintenance branch in the US
- Dallas, TX - Strong rental/lease as well as maintenance

The branches selected were also chosen because of their strong leadership. "We knew they would be open to trying the new process, and work with us if any bugs had arisen" said Gianetti. They were also relatively close to home base for many of the team members. We each had a branch to pick apart.

Purchasing was the first part of the implementation. Since the branches have been using the PI system for buying other parts, it was a natural progression to start ordering tires. Training had to be conducted on how to read delivery receipts, and tires DOT information. The "cheat sheet" was also given out to aid in the training process. The branches were given two weeks to get their ordering process straightened out. The ordering also produced some "bugs", which were more process breakdowns than IT "bugs". The branch would have to call or send a fax to the tire dealer listing how many tires they needed in a given size. The dealer would then send tires to the branch with the delivery receipt detailing what "went into the tire". If a branch did not enter this information quickly enough the tire dealer would send an invoice with no PO number listed. This would delay payment.

PI had a functionality to have a "shopping cart" preloaded with all tire options. A list was also made to have the branches quickly add up any repairs or other items. Instead of a branch adding a line for every tire, they were told to add a line for every piece of work performed with various quantities. For example, a branch orders five tires, they enter a line for casings with a qty of five, three nail repairs, five treads and five valve stems.

Once the branches were using the purchasing part effectively the team moved on to inventory. The inventory process required more training, since this was completely new to the branches; they had never used this function in PI.

Since the inventory hardware and software were in place at the branches the team members only had to "flip the switch". Conducting the branch inventory was an easy task; tires were not recorded by size, brand, and tread design. Only the on hand replacement inventory was counted. The branch staff did the counting under supervision and direction of the team members. The information on casings at the vendor location was taken from the "old" tire inventory log spreadsheet.

One person from each branch was to be named the "tire person". While all staff was trained on the use, one person would be the local expert, and held accountable for branch compliance and performance. The tire person would be responsible for weekly cycle counting and balancing of the inventory. They would also be responsible for ensuring national account usage, and proper documentation of use, and management of the program. Tires had to meet the TFS specification or should be rejected.

Branch management also had the autonomy to change the spec given certain conditions (with approval). Storage trailers, which did not travel on the road too much, are generally, left at a site for storage purposes. These typically older trailers did not need new first quality tires, so the branch could use their older tires avoiding expense. The same could be said for tires on equipment that was held for sale. Equipment is sold as is; so older tires were used since new ones would not increase the resale value. Inventory management only helped this process because the branch could look in PI and know exactly what they had on hand. This saved time and money.

The biggest training effort was with the mechanics. They normally would not add actual parts to their repair orders. In the past branch staff entered an estimated cost; some "strong" branches would use actual pricing and add everything up. This was very time consuming. The new process would take the average tire cost and automatically add it to a repair order (RO). The mechanics had to be trained on adding the parts. Branch management had to be trained on resolving discrepancies. If a part was added to a RO and the part was not in the system, it would be rejected, and the RO could not be closed until the inventory discrepancy was resolved.

This had several major impacts. Branch management could review their RO's and find errors. The system would tell them if a mechanic took it upon himself to buy a part off the national contract, another point of accountability. They were also able to see which tires were moving fast, and which were not. The system also has the built-in functionality to let the ordering people know when an item is getting low in inventory. The limits initially set by the corporate staff could be adjusted locally.

The piloting and implementation has expanded, but is not yet live in all locations. It has caused dramatic improvement in branch productivity. At first most of the detractors did not want to learn the new process. Once they became accustomed to it and had some practice, they could not live without it. The branches that have not been implemented yet are even clamoring for it after hearing about its many positive impacts from their colleagues.

This project was a simple change with limited IT investments, building upon a system that had already been in place. It has potential for tremendous ROI. Before, the analytics were available most customers were charged a flat rate per mile. Now, with the visibility TFS can increase the price and margins for the poorly performing customers, and lower the cost for the customers that have strong maintenance programs in place. It has also been a value add, because we can show our customers how to be more effective in several areas of their business.

For a trucking company fuel and downtime are their CTQs. By having an effective tire management program in place we can limit both. The savings not only impact our business, but also have actually increased Net Promoter Scores in the branches that have had the process implemented. This leads to more repeat customers, as well as great word of mouth. In an industry as tight knit as trucking, this has been a huge benefit.

TFS is even considering the tire management as a new product in its Total Fleet Management strategy. This would take maintenance and management out of the customer's hands allowing them to focus on their core competencies. Implementation should be completed in the third quarter, and all US branches will be onboard. It is estimated product savings will be around $1.2MM. The soft savings are much more difficult to quantify.

REFERENCES

References available upon request to the authors

RELATIONSHIP AMONG ATTENDANCE, CLASS SIZE AND EXAM GRADES IN THE INTRODUCTORY BUSINESS STATISTICS COURSE

Boualem Kezim, Merrimack College, 315 Turnpike Street, North Andover, MA 01845, kezimb@merrimack.edu, (978) 837-3454
Susan E. Pariseau, Merrimack College, (978) 837-5417, susan.pariseau@merrimack.edu

INTRODUCTION

Many individuals believe that both attendance and class size are key factors in student performance in college-level courses. We examine these beliefs in the introductory business statistics course at two institutions of higher education. Since statistics is considered to be a challenging course with an emphasis on quantitative material, we believe that poor attendance will negatively impact performance in the course and that there will be a nonlinear relationship between these two variables. We believe that the performance of students who miss a small number of classes will not be greatly impacted while the decrease in performance will be magnified as the number of absences increases. We also believe that, in general, if students are placed in smaller classes, they will perform better than if they had been placed in a large class. In small class environments where students tend to know their classmates, we believe that, at least to some extent, absences will not have as great an impact on performance.

LITERATURE REVIEW

A review of the literature finds multiple articles reporting the effects of absenteeism on student performance in college. Romer (1993) reported that roughly one third of students at elite American schools are absent from their economics courses on any given day. Many other authors have studied the effects of absenteeism on students' performance in economics courses (Park and Kerr, 1999; Durden and Ellis, 1995; Marburger, 2001; Glasure, 2002; Krohn and O'Connor, 2005). Using a variety of performance measures and variables (including ACT, SAT, GPA, course and exam grades), several studies addressed the impact of absenteeism on students' performance in selected business courses including introduction to business (Patron and Bisping, 2006), operations management (Peters et al., 2002) and principles of finance (Chan et al., 1997). While Park and Kerr (1990), Romer (1993), Durbin and Ellis (1995), Chan et al (1997), Marburger (2001), Glasure (2002) and Cheung and Kan (2002) report that attendance positively affects student performance, other studies found conflicting evidence between class attendance and student performance. Peters, Kethley and Billington (2002) found that for operations management students with graded homework, there was no significant correlation between attendance and total exam points. When homework was not graded, however, a significant correlation was found. Krohn and O'Connor (2005) found a positive relationship between attendance and course average in intermediate macroeconomics courses but no relationship between attendance and scores on individual examinations.

While many authors focused on the impact of attendance on student performance in various business and economics courses, others limited their study to the impact of the class size component on student performance. There has been a substantial amount of research on how student performance is affected by class-size at the college level and the primary and secondary levels [see Fenollar et al (2007)]. At the elementary level, Nye and Hedges (2001) reported that there were cumulative positive effects in reading and mathematics for smaller classes. At the high school level, Rice (1999) found that class size appeared to have an effect on performance in mathematics classes but not in science classes. Focusing on the influence of high school science class size on students' achievement in introductory college science courses, Wyss et al (2007) showed no difference in student achievement until class sizes were reduced to ten. At the college level, Hancock (1996) reported no significant difference in performance among students in different size sections of statistics classes. This study did not consider any factor other than class size. In an introductory marketing course, Dommeyer (1997) found that overall, students from large classes performed as well as students from small classes on the graded component of the course. In contrast, Raimondo et al (1990) showed that the large lecture has a negative effect on learning in an introductory economics course. Likewise, Arias and Walker (2004) reported that small class size has a positive impact on student performance in the principles of economics course. Raimondo et al (1990), Arias and Walker (2004) and Wyss (2007) used multiple regression models.

This paper investigates the combination of two factors, attendance and class size, on performance in an introductory business statistics course.

METHODOLOGY

This study was conducted at two institutions of higher education in the northeast region of the United States, a small private undergraduate college and a large private university. Our major goal is to study the effects of attendance and class size on student learning in business statistics. In the small college, data were collected from five small sections of a required Business Statistics course; at the large university, data were collected from one large section. The majority of the students enrolled in these courses were sophomores. Students in all classes were graded based upon their performance on two exams, a comprehensive final exam, collected homework, class participation, and class attendance. Both the large section and all the small sections of business statistics were taught by the same professor using the same teaching methodology. Across the five small sections of the Business Statistics course, all groups had identical exams and homework and used the same textbook. In the single large section, a different textbook was used but the exams were similar to the ones given in the small sections.

In general, statistics is viewed as a hard course among business students. Due to its quantitative aspect, we believe that attendance plays a key role in explaining students' performance. In our analysis, we also included the students' GPA, another factor that is commonly alleged to affect student performance.

In the small college statistics course (small class), class attendance was recorded at every meeting while in the large university (large class), attendance was recorded randomly at approximately fifty percent of the class meetings. In the small course, combined enrollment of the five sections resulted in a total sample of 110 students excluding students who dropped the course. Eighty-six (86) students were enrolled in the large course.

In both courses, class attendance was worth five percent of the final grade. The median percentage of absences was 12.5 and 14.3 for the small and large classes respectively. During the semester, we observed that almost 19 percent of students in the large class missed more than 20% of the lectures whereas in the small class only 9 percent missed in excess of 20%. Overall, the absenteeism rate was higher for the large course. The mean absenteeism was 15.1 percent for students enrolled in the small class versus 20.2 percent for those enrolled in the large class. The mean GPAs were comparable for the small class (2.91) and the large class (3.04).

ANALYSIS

Table 1 shows the Pearson correlation coefficients between two factors considered to be related to students' performance in a business statistics course. To measure students' performance, we used only the weighted average scores of the examinations given in class and excluded homework, class participation, and attendance.

Table 1. Pearson Correlation

ExamAvg	GPA	Abs%
Small Class	0.714**	-0.446**
Large Class	0.546**	-0.588**

** Correlation is significant at the 0.01 level (2-tailed).

As expected, table 1 suggests that the variable *ExamAvg*, which measures student performance in the business statistics course, is significantly negatively correlated with absenteeism and positively with GPA for both class sizes. We also found that *GPA* and *Absence Percent* are highly negatively correlated (-0.418**). This may be explained by the assertion that students with high GPAs have a tendency to miss fewer classes than students with low GPAs. In order to further explore the relationship between absences and academic performance, these two factors, *Absence Percent* and *GPA,* were tested using a multiple regression analysis to measure students' performance in a business statistics course. Both variables were found significant.

Although absence percent is highly correlated with performance, we hypothesize that there is a nonlinear relationship between absences and performance as reported by Durden and Ellis (1995). Few absences have little impact on performance. As absences increase, the effect on performance increases rapidly in a nonlinear manner. Thus we include the variable absence percent square (*Abs%Sq*) in our model to depict this effect.
We included an indicator variable (*Large Class*) for class size (1 if large class and 0 otherwise) and an interaction term between *Absence Percent Square* and *Large Class*

(*IntSq*) in our model below. This interaction variable will allow us to determine whether there is a difference in the magnitude of the effect of absenteeism on students' performance between large and small classes. Although various multiple regression models were considered and tested for goodness-of-fit, the following multiple regression model was retained:

$$ExamAvg \;=\; \beta_0 + \beta_1\, GPA + \beta_2\, Abs\%Sq + \beta_3 L\arg e\, Class + \beta_4 IntSq + Random\, Error$$

Table 2 reports the ordinary least-squares estimates with *ExamAvg* as the measure of student performance and *GPA*, *AbsencePercentSquare*, *LargeClass* and *IntSq* as the explanatory variables.

Table 2. Least Squares Estimates

Variable	Coefficient Estimate	t-statistic	P-value
Constant	41.77	9.25	.0000
GPA	12.87	8.77	.0000
Abs%Sq	-27.63	-2.65	.0009
LargeClass	5.33	3.25	.0013
IntSq	-28.71	-1.92	.0565

R square = 0.49 ; R square Adj = .48 ; F Statistic = 46.15 (.000)

The results of our regression analysis yield the following fitted model:

$$ExamAvg = 41.77 + 12.87 GPA - 27.63 Abs\%Sq + 5.33 LargeClass - 28.71 IntSq$$

Table 2 reveals that *GPA* is the most significant predictor of students' performance. The *GPA* least square estimate indicates that for every additional point in the cumulative GPA, the average exam grade is expected to increase by 12.87 points. The indicator variable coefficient (*LargeClass*) shows that the exam grades of the students in the large class are, on average, 5 points higher. The interaction factor reveals that absences have a greater impact in a large class setting than a small class environment. A nonlinear relationship between absenteeism and performance is present in our model meaning that a small number of absences will have little impact on student performance while an increasing number of absences results in a quadratic decrease in average exam grades.

Looking at the coefficient estimates of the predictor *Absence Percent Square* in the two fitted models below, it is clear that absences result in greater decreases in grades for students in large classes. The *Absence Percent Square* coefficient for the *Large Class* is double the size of the coefficient for the *Small Class*.

Small Class: E (Exam Average) = 41.77+12.87*GPA* – 27.63 *Absence Percent Square*
Large Class: E (Exam Average) = 47.10+12.87*GPA* – 56.34 *Absence Percent Square*

Table 3. Effect of Absences on Exam Average for Small and Large Classes

% Absence	Small	Large
10	-0.276	-0.563
20	-1.105	-2.253
30	-2.486	-5.070
40	-4.420	-9.013
50	-6.906	-14.083
70	-13.537	-27.603
100	-27.626	-56.333

Figure 1. Effect of Absences on Exam Average for Small and Large Classes

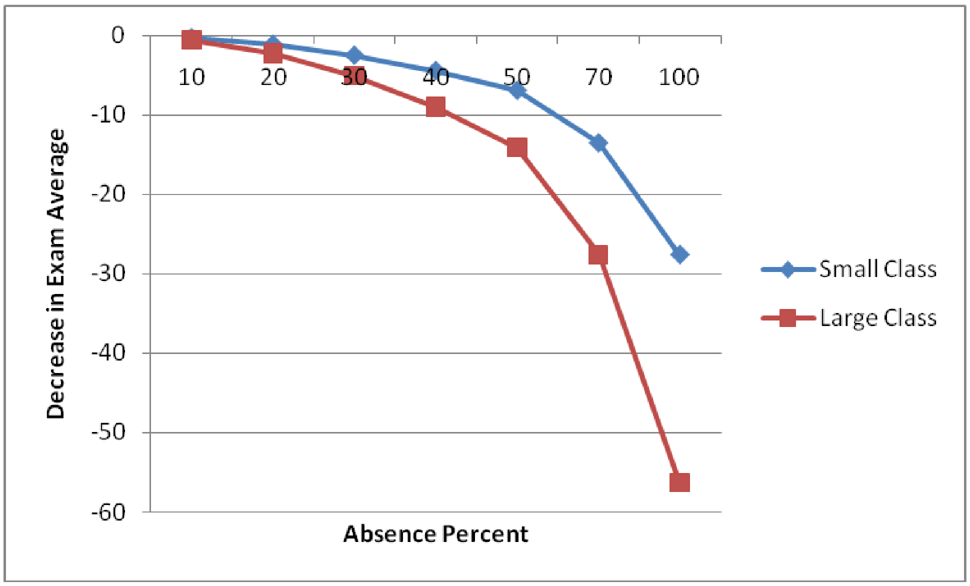


Table 3 and Figure 1 clearly illustrate that a small percent of absences will not significantly affect student performance in either large or small classes. However, an appreciable amount of absences will produce a negative quadratic impact on student performance. The fitted models, Table 3 and Figure 1 all provide a comparison of the effects of absences on exam average for small and large classes.

CONCLUSIONS AND LIMITATIONS

Similar to Patron & Bisping, Park & Kerr, Romer and Durden & Ellis, we found a statistically significant positive relationship between performance and GPA. In our model, GPA was the best predictor of student performance.

We found a similar relationship between absenteeism and performance to that reported by Durden and Ellis. These authors incorporated multiple indicator variables to represent categories of total absences (e.g., 1 or 2 absences; 3 or 4 absences, etc.) whereas we used a single quadratic explanatory variable expressed as the square of the percent absences. Since our variable squared the percent absences, it captured the fact that as the number of absences gets larger, the effect on student performance, measured by weighted average exam grade, increases quadratically. Similar to our study, Durden and Ellis did not report significance until the number of absences was at least five. We concur, therefore, that a small number of absences has little effect on student performance. When a certain threshold is reached, however, the effect is noticeable and increased absenteeism above this threshold magnifies the impact on performance.

Similar to Raimondo et al (1990) and Arias and Walker (2004) in their study of introductory economics, we found that the large lecture has a negative effect on learning in business statistics. Our findings demonstrated that absences have a greater impact on student performance in a large class. We found, for example, that a student in a small class who has missed fifty percent of the classes will have approximately a 7 point lower exam average than one who has missed no classes. A student in a large class, with the same number of absences, would reduce their exam average by 14 points.

Our research has several limitations. The students are drawn from different populations. Although both institutions are selective in their admissions processes, one is considered highly selective. Although the same professor taught all classes using the same pedagogy, different texts and exams were used. Attendance was taken at every class meeting of the small class but only fifty percent of the time in the large class. These limitations may have had a confounding effect on the results provided in this paper. Ideally, the same professor, using the same exams and texts, will teach small and large classes at the same institution with random placement of students in the classes of varying size. In reality, this situation is difficult to achieve.

REFERENCES

References are available from Boualem Kezim upon request.

USING EXCEL IN AN INTRODUCTORY STATISTICS COURSE: A COMPARISON OF INSTRUCTOR AND STUDENT PERSPECTIVES

Cynthia L. Knott
Marymount University, 2807 N. Glebe Road, Arlington, VA 22207
cynthia.knott@marymount.edu
703-284-5727

G. Steube
Marymount University, 2807 N. Glebe Road, Arlington, VA 22207
gsteube@marymount.edu
703-284-5943

Northeast Decision Science Institute

March 2010

Abstract

Almost all undergraduate business degree programs require that students take at least one course in statistics. The instructor for these classes has a number of options in terms of selecting the software package that will be used. Three popular options are R, PASW (formerly SPSS), and Excel. This paper examines the advantages and disadvantages of these software choices in order to assist the instructor with making an informed choice. The strengths and weaknesses of these choices are explored in terms of both the instructor's and student's perspective. After reviewing these assessments, the paper concludes that Excel is the best choice because of its low cost, wide availability, familiar interface, and computational and charting flexibility.

Although most business schools require at least one course in statistics advances in technology and the availability of statistical software has dramatically changed the way in which these classes are taught. In the past, instructors presented statistics the old fashioned way, with the "paper and pencil". The students were expected to be involved in the mathematics as well as the calculations. Students learned how to use the z-tables and calculate the statistics by hand. Because of advances in technology, the reduction of cost in acquiring technology, and the ubiquitous presence of technology, the way in which statistics courses can be presented has changed to offer the instructor a wider range of choices. These choices center on the software package that will be used in the class to perform calculations and present the findings in graphs and charts. This paper focuses on some of the choices that instructors may face in selecting the best software support for their introductory statistics class.

In the business school, statistics courses are applied rather than theoretical. It is important that these learners understand the underlying calculations and theories, but more significantly, these students must be able to apply, analyze, and interpret the results to improve decision making in the business environment. Communication is the top skill sought by most employers today (Barnes, 2009). Statistics courses should help the student in learning how to communicate the results of their analysis in their future business environment. Therefore the business statistics courses emphasize the communication of the results obtained from conducting a wide range of hypothesis testing and other types of analysis. The use of statistical software relieves the student and instructor from spending too much time on calculations and thus affords more time to emphasize the understanding, interpreting, and communicating the results of these computations.

This paper compares three popular statistical packages; R, PASW (formerly known as SPSS) and Excel. Information about the R package can be obtained from the *R Package for Statistical Computing* (Department of Mathematics and Statistics at Vienna University, n.d.); PASW and Excel information can be located on the PASW website (SPSS Inc., 2009) and the Microsoft website (Microsoft Inc., 2009b). The use of any of these three packages has a number of strengths and deficiencies. These advantages and disadvantages are discussed from the instructor and student point of view in this report. The discussion will provide input by which assist an instructor in deciding among the three packages for his or her class in introductory statistics.

Advantages and Disadvantages of R

The advantages and disadvantages from the instructor and student views for R are summarized in Table 1 (instructor) and Table 2 (student), respectively.

Table 1

Advantages and disadvantages of R from the instructor view

Advantages	
1. Availability	Freely widely available at no cost (Zieffler & Long, n.d.)
2. Flexibility and customization	Because R is a programming language almost any result can be achieved (Zieffler & Long, n.d.)
3.Up to date methods and packages	Because R methods are written by users, R is more current than many commercial statistical packages that require updates to their base system (Zieffler & Long, n.d.)
4. Broad coverage	R packages are extensive and include a wide variety of quantitative applications
5. Availability of help	Because R has a large user network, help is readily available on almost any topic

Disadvantages	
1. Large data sets	R may not handle large data sets as efficiently as SAS (Zieffler & Long, n.d.)
2. Speed	Some procedures in R could take days to run (Zieffler & Long, n.d.)
3. Learning curve	Because R is command line driven rather than a point and click application, the learning curve is more challenging than most commercial packages which use a graphical interface
4. Lack of a spreadsheet view of data	Unlike Excel and SPSS, R does not include a spreadsheet view of the data set
5. Unfamiliarity	R is less well-known than SPSS or Excel and may be viewed less positively because of the lack of familiarity

Table 2

Advantages and disadvantages of R from the student view

Advantages	
1. Availability	No cost to the students and can be installed on their computer and thus eliminate trips the University's computer lab to do homework
2. Flexibility and customization	After some investment in learning, R the student could find available packages and develop their own tailored applications
3.Up to date methods and packages	As other users add packages to R to keep it current, students can continue using R throughout their academic and professional careers without concern about updating their base packages to obtain the newer software
4. Broad coverage	Because of R broad coverage, students might be able to use R in other quantitative courses
5. Availability of help	Students can use the network of available user without charge or other registration requirements to obtain guidance on almost any aspect of R

Disadvantages	
1. Large data sets	May not be an issue for a student's academic use of R because the data sets are relatively small
2. Speed	The basic procedures used in an introductory statistics course would not take days to run
3. Learning curve	Students would need to invest time in learning R and that investment may be challenging and time intensive; the lack of graphic interface with point and click capability would be a challenge for students who have only worked with operating systems and applications that furnish these abilities
4. Lack of a spreadsheet view of data	The lack of spreadsheet view may be a concern to students because they are tuned to table presentation of data especially for business students
5. Unfamiliarity	Students may not be any more unfamiliar with R than SPSS; most students would be familiar with Excel

Advantages and Disadvantages of PASW

The strengths and weaknesses from the instructor and student views for PASW are summarized in Table 3 (instructor) and Table 4 (student), respectively.

Table 3

Advantages and disadvantages of PASW from the instructor view

Advantages	
1. Well known and supported	PASW has been in the marketplace for many years and many textbooks for introductory statistics courses are based on this application. Other programs can easily import SPSS data files (Harrington, McLeod, & M.Clark, 2009).
2. Ease of use	Because of its graphical user interface, a large of number of statistical functions are easy to use and access
3.Up to date	PASW is generally updated at least once a year with minor updates available on the corporate website

4. Broad coverage	PASW includes a large number of statistical routines appropriate for an introductory statistics course
5. Availability of help	There are number of Excellent books that provide comprehensive information on using PAWS; the availability of the list serve for this package also provides another source of guidance for specific issues
Disadvantages	
1. Cost	PASW is a commercial software that is available at a relatively high cost [JourneyEd (2009) list PASW Graduate Pack for Windows at $199.98] ; some applications are available only as another product with a separate fee
2. Licensing complexity	The license to use PASW is time limited; the license only allows installation on a limited number of computers; add-ons to the PASW require acquisition to additional licenses. The licensing is not user friendly (Harrington, et al., 2009).
3. Confusion among the different versions	Because PASW is updated every year a number of issues can occur with features and data formats are available especially between the MAC and Windows products

Table 4

Advantages and disadvantages of PASW from the student view

Advantages	
1. Well known and supported	Students would have relatively little difficulty in obtaining support for their SPSS work from textbooks, Internet resources, and other students.
2. Ease of use	The graphical user interface provides the student with quick access to the statistical routines needed for an introductory course; student also has the ability to customize PASW results including the display of graphic output
3. Broad coverage	All the statistical routines required by the introductory

	statistics course would be available to the students
4. Availability of help	The student has access to a large number of excellent books that provide comprehensive information on using PASW
Disadvantages	
1. Cost	At price point of almost $200 even for the Graduate version of the PASW package, the cost is an issue for student
2. Licensing complexity	Although students would be able to use their purchased package for the duration of the introductory statistics course subsequent use of the package in graduate and professional work would become a licensing problem
3. Confusion among the different versions	Students may try to purchase or borrow older versions of PASW which could create issues for them with the instructions in the selected introductory statistics textbook
4. Lag in newer techniques	"For academic use SPSS lags notably behind SAS, R and even perhaps others that are on the more mathematical rather than statistical side for modern data analysis (e.g. robust and bootstrapping approaches available easily conducted elsewhere are nonexistent or very difficult to do, basic tests of analytical assumptions are often not available)" (Harrington, et al., 2009)

Advantages and Disadvantages of Excel

The advantages and disadvantages for Excel are provided in Table 5 (instructor) and Table 6 (student).

Table 5

Advantages and disadvantages of Excel from the instructor view

Advantages	

1. Availability	All of the labs in the University have the Microsoft Office Package on them and therefore, access is always available. Instructors don't have to go through the hassle of having "special" software installed on the computers each semester.
2. Availability of Help	On-line help is extensive and the package also has a built in help function that is very user friendly.
3. Ease of Use	Very hands on friendly and intuitive to use. The dropdown menus make it easy to find things.
4. Add-Ins	The software has a standard installation, but you can also add-in data analysis tool packages, specifically to do Statistical applications. Therefore, you don't have to know any coding or programming languages.
5. Coverage	The package includes all of the Statistical applications that an introductory course needs.
6. Data Sets	Many authors are including data sets with the textbooks that are already in Excel files and ready for analysis. Also, many web sites that manage data sets are putting them in a format that is easy to download as an Excel file.

Disadvantages	
1. Cost	Although the package is available in all of the labs on campus, an instructor would need to purchase the software package for any home computers that they use.
2. Potential Calculation Problems	The program doesn't calculate the 3rd quartile correctly (Anderson, 2009). Problems with the calculations in Excel have also been identified by McCullough and Heiser (2008) and Yalta (2008).
3. Limitations of Data	The number of data points is limited to x amount. In terms of an introductory course this is not necessarily an issue, but it can be when showing analysis of large data sets. The limitations of Excel are provided by

	Microsoft (2009a).
4. Functions	Some of the advanced statistical functions are not included in the package.

Table 6

Advantages and disadvantages of Microsoft's Excel from the student view

Advantages	
1. Availability	All of the University labs have Microsoft Office installed on them, so the students have access to Excel everywhere on campus.
2. Cost	There is no cost to the students if they use the computers in the lab.
3. Availability of Help	There is extensive on-line help and the package also includes a help function that is easy to find information.
4. Textbook	Many of the textbooks include instruction on the functions in Excel, which allows students to follow along and do practice problems on their own.
5. Data files	Many of the textbooks are including data files that are already in Excel format and also many web sites that manage data sets are making the files available for download in Excel format.
6. Applicability to other Courses	The use of Excel is becoming the standard in many other areas of business instruction such as accounting, finance and operations. Therefore, the students can use the skills they learn in other courses as they work through their programs.

Disadvantages	
1. Customization	The ability to customize the package is not very user

	friendly. You can add macros using visual basic, but this requires the knowledge of a computer programming language.
2. Large Data Sets	The limitations of Excel are provided by Microsoft (2009a).
3. Confusion among the different versions	Because Excel is in a current transition from 2003 to 2007, students sometimes are working in one version and being taught in another; this can be confusing to them. Although the functions are all the same, where to find them is different for each version.

Conclusion

This paper explored three popular software packages to determine their advantages and disadvantages for use in an introductory business statistics course. By assessing these pros and cons in terms of both the instructor and student perspective a better decision about which of these package to choose can be made. Although any of the three packages would provide learners with the ability to perform computations and display graphs and charts, each of these choices has its own unique set of strengths and weaknesses. This report identified a number of assets and liabilities for each of the three packages in Tables 1 thru 6. Because of its low cost, computational and charting flexibility, interface familiarity, and wide availability the selection of Excel is seen as the best choice. The availability factor for Excel includes not only its presence in college computer labs and classrooms but also its wide use in businesses. After students complete their business degrees and are employed in their fields, the availability of Excel in these future workplaces will far exceed SPSS or R. By selecting Excel for use in the statistics course, the instructor has added to the motivation for his or her students because in almost all of their potential employment destinations, the former student will be able to use Excel on the job for their data analysis needs. Consequently Excel is an excellent choice for instructors to use to support their introductory statistics class.

References

Anderson, D. R. (2009). *Essentials of modern business statistics* (4th Ed. ed.). Eagan, MN: Cengage South-Western.

Barnes, K. (2009). Skills Most Sought After by Employers Retrieved October 24, 2009, from http://iccweb.ucdavis.edu/LAB/articles/Skills.htm

Department of Mathematics and Statistics at Vienna University. (n.d.). The R package for statistical computing Retrieved October 24, 2009, from http://www.r-project.org/

Harrington, R., McLeod, P., & M.Clark. (2009). SPSS short course. Retrieved October 12, 2009, from http://www.unt.edu/rss/class/SPSS/course1.htm#VI.__Relative_Advantages_and_Disadvanatges_

Journey Education Marketing, I. (2009). PASW Statistics Grad Pack 18.0 for Windows (Formerly SPSS). Retrieved October 12, 2009, from http://www.journeyed.com/item/SPSS/PASW+Statistics+Graduate+Pack/100966436?sSRCCODE=GOOGLEBASE&utm_source=GoogleBase&utm_medium=Comp%2BEngines&utm_campaign=jce01

McCullough, B. D., & Heiser, D. A. (2008). On the accuracy of statistical procedures in Microsoft Excel 2007. *Computational Statistics and Data Analysis, 52*(10), 4570-4578.

Microsoft Inc. (2009a). Excel specifications and limits Retrieved October 26, 2009, from http://office.microsoft.com/en-us/Excel/HP100738491033.aspx

Microsoft Inc. (2009b). Microsoft Office Excel 2007 Retrieved October 25, 2009, from http://office.microsoft.com/en-us/Excel/FX100487621033.aspx

SPSS Inc. (2009). SPSS an IBM company Retrieved October 24, 2009, from www.spss.com

Yalta, A. T. (2008). The accuracy of statistical distributions in Microsoft Excel 2007. *Computational Statistics & Data Analysis, 52*(10), 4579-4586.

Zieffler, A., & Long, J. D. (n.d.). Basics of R. Retrieved October 12, 2009, from http://docs.google.com/gview?a=v&q=cache%3AKNt87a_hJBAJ%3Awww.tc.umn.edu%2F%7Ezief0002%2FNotes%2FRBasics.pdf+advantages+of+R&hl=en&gl=us&sig=AFQjCNFNcPf8U4xtjQ3vC-qvzUcCqYod0w&pli=1

An Exploratory Study Investigating Postgraduate Students studying for the Chartered Institute of Marketing Diploma in Marketing and their Preferred Modes of Learning and Study

Keywords: Education, Learning, Teaching, Case Study, Knowledge Transfer

Abstract

The authors have been involved in teaching on the UK professional qualifications for the Chartered Institute of Marketing (CIM) Postgraduate Diploma (now re-titled the Professional Postgraduate Diploma) Analysis and Decision (case study) paper. The authors also teach on postgraduate masters courses and case studies are used in these programs as well. The pressures faced by mature, professional students in maintaining a balance between work, home and study prompted the authors to undertake a study to ascertain whether the teaching methods employed on the case study are the most efficacious in terms of learning styles and work-life balance.

The case study requires a great deal of student participation and group work and not all students enjoy, or benefit from, this approach. Those who do not engage in-group work tend to experience a higher risk of failure in the examination.

Teamwork, interaction and peer knowledge transfer are all-important aspects of success. Those who benefit most are able to link theory with experience to provide comprehensive analysis and solutions to case problems.

The data highlights that evenings only, or afternoons and evenings, are an acceptable mode of study to the cohort. On line/distance learning is also shown to be an

acceptable mode. However, the most preferred mode of study is evenings only. Students' least preferred modes of learning were intensive courses, online or distance study and in-company courses.

Introduction

The University Business School is committed to explore areas where it can improve the quality education and student learning experience provided to its customers. The authors have a combined 45 years of working in industry and higher education and thus recognise the pressures placed on professional students and the need to maintain a balance between personal life, career and professional study.

Objectives

The aim of the research is to explore the general issue of learning styles, modes of study, modes of learning, the quality of post graduate teaching, external pressures on students and then to specifically look at group work in relation to the analysis and decision case study.

The specific objectives are to:

- Analyse the learning preferences of CIM case study students
- Understand those preferences in the context of their individual learning styles, the CIM course, the case study as a mode of learning, and their work-life commitments
- Relate these outcomes to current educational theory and practice

The Case Study Method

The case study is a practical test of the student's knowledge (gained in the CIM's Certificate (Professional Certificate in Marketing), Advanced Certificate (Professional

Diploma in Marketing) and Post Graduate Diploma (Professional Post Graduate Diploma in Marketing) or equivalent studies of marketing and their ability to apply it. Normally students will also have some practical experience of marketing to bring to bear.

Case studies can ask you to do a number of things:

- Understand a situation and its implications
- Solve specific problems

The CIM Analysis and Decision case study paper contains the following characteristics:

- Lack of a formal syllabus; students cover theory in other modules
- A student-centred teaching approach with little structured tuition
- The application of skills such as analysis and synthesis
- An emphasis on team work and group dynamics in the preparation phase, although students prepare their own marketing plans for the final three hour examination in which they have to answer three questions based on the case
- The application of theory learnt in other CIM Diploma modules to an extensive real-world situation

Students are encouraged to conduct in-depth discussion with colleagues on the case study analysis and its issues. This is accomplished through syndicates of approximately six students and the holding of frequent plenary sessions where the complete group gather together. In this way a syndicate member will not only hear the views of his or her syndicate but also those of other syndicates. By working in this manner a more balanced, integrated and secure approach can be developed.

Ethical Considerations

The research will take into account the following ethical issues:

- Anonymity of students
- Agreement to use quotes from students in the final research report
- Offer students the opportunity to receive a copy of the research

Literature Review

A case study is the detailed description of a particular real life situation or problem as it happened in the past or as it could happen in the professional life of the student (Kreber, 2001).

It is well known that after three decades of experimenting with learning styles the effectiveness of first identifying and then complementing how each person begins to concentrate on, process, internalise and retain new and difficult information and skills is proven (Dollar, 2001).

The potential value of the case study method of teaching has been demonstrated by leading practitioners such as the London Business School and Harvard Business School (Christensen & Hansen, 1987, cited in Kreber, 2001).

Gross Davis (1993) summarises research into good case study practice as:

- A good case study tells a story with a thought-provoking issue
- Contains elements of conflict which promotes empathy with the main characters
- Lacks an obvious straight-forward answer or solution
- Requires students to develop a strategy and make a decision

- Group work is important in case study analysis and development

Whilst learning styles are an important basis on which to develop teaching strategies, students need to become actively involved in case study exercises. In order for these to be a valuable learning tool students must commit to attending regularly and be prepared to discuss and exchange ideas in class Shapiro (1984). It is important to recognise that different students and their respective learning styles will lead some to participate whilst others may feel uncomfortable in that type of environment.

Despite a variety of excellent books on alternative teaching strategies in higher education, for example Cranton (1998), Gross Davis (1993), McKeachie (1986), university teaching tends to emphasise abstract concepts and theories rather than practical or concrete experience. Furthermore, the encouragement of internal reflection (accompanied by a reflective log) as a way of capturing the meaning of students' experiences are fairly limited in use (Kreber, 2001).

Concrete experience combined with abstract conceptualisation (the application of appropriate theories and models) need to be combined for learning to become experiential and, ultimately, valid as a means of learning. Kreber (2001) describes learning as being 'transformational' and case studies combine concrete and abstract events into knowledge.

The output from classroom based case study work depends strictly upon student input and how active a role they are prepared to play. According to Leenders, Erskine and Leenders (1997, p7) to achieve positive outcomes students must:

- Help by teaching others through plenary sessions
- Actively participate

➢ Take risks

➢ Learn from instructors and classmates.

This is further supported by Kolb's (1979) (Figure 1) views regarding experiential learning e.g. the participation and risk taking and then reflecting on experiences and improving performance from that.

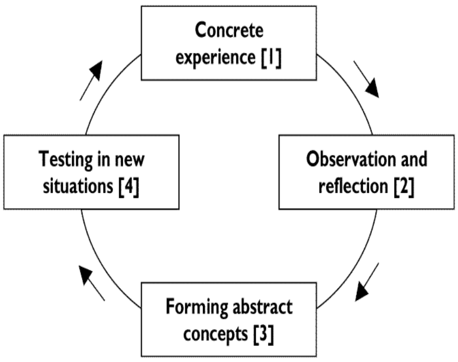


Figure 1: Kolb's Experiential Learning Cycle

Source: http://www.infed.org/biblio/b-explrn.htm

A case study, properly facilitated, is an effective way of enabling students to become involved in all phases of Kolb's cycle and will develop students' higher-order learning such as critical thinking ability (Kreber, 2001). It will also develop students' propensity for self-learning (Guglielmino, 1977, cited in Kreber, 2005).

Kolb (1984), building on the work of Dewey (1938) and Lewin (1951) also emphasises the importance of reflection in building on experience and hence meaningful learning as well as the cyclical nature of the learning process (cited in Kreber, 2001).

Students who use both comprehension and operationally-based learning techniques are most likely to gain the deepest level of understanding (Cuthbert, 2005).

If we consider the three stage learning process presented by Leenders, Erskine and Leenders (1997, p19), (Figure 2) then perhaps we can see why participation is so important to obtain full benefit from the case study process. Some students prefer to work on their own and fail to gain benefit because of their reluctance to participate in small and/or large group discussions owing to their respective learning style. This would imply that the greater the participation the greater the quality and quantity of input and output.

On the other hand:

> **Students that forego plenary sessions and group discussions are likely to fail, no matter how individually clever at analysis they might be.** ***Pearson (pp vii, 1995).***

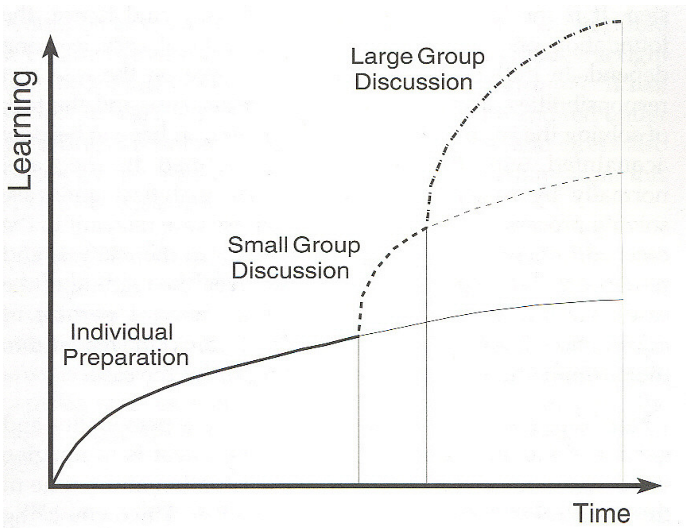


Figure 2: The Three Stage Learning Process

Leenders, Erskine and Leenders

Success in terms of large or small group discussions is only really likely to provide positive results depending upon the characteristics and cohesiveness of the group (Greenberg and Baron, 1995).

The other influence upon the effectiveness of the group is whether they perceive themselves as being either formal or informal. Formal is based around task-orientated activities whilst informal tends to be friendship/interest groups (Greenberg and Baron 1995). The nature of the CIM case study analysis will lead the group into a formal role although it is quite possible that a mix of both will be experienced by the members. Whilst the CIM case study has no formal syllabus, analysis of the case still involves a learning process facilitated by teaching staff.

Student activities and engagement such as active involvement with peers and sharing the experience of learning (in this instance via case study work) have been found to enable students to develop a network of support (Tinto, 1997) in addition to being an important component of academic success (Finn and Rock, 1997). Group integration and collaborative activities reinforce the learning process Terenzini (1999).

Work-life balance is important to professional people, particularly when they have to study for further examinations in their own time. Research by Tausig and Fenwick (2001) suggests that perceived control of work schedules increases work-life balance net of family and work characteristics. Control thus becomes an important element in work-life balance. Having the flexibility and time for further study is one element in this equation.

Methodology

Questionnaires were distributed to 60 current and ex-SBS students who had studied since October 2000. Twenty-eight (28) completed questionnaires were returned to support the research. Following a trial to detect potential errors in question design and wording the questionnaires were emailed for self-completion or completed during classroom sessions, following a trial to detect potential errors in question design and wording.

Structured questions utilise the Likert scale technique enabled respondents to express degrees of relative preference (Easterby-Smith et al, 2002). Respondents were also asked to rank preferred learning modes across the range of CIM modules studied.

The data was cross-referenced with an earlier study of CIM Wessex branch members, which examined preferred learning styles for CPD activities.

Reliability of Data

The reliability of any questionnaire is always an issue for researchers. Much has been written about questionnaire design and analysis of results.

Saunders, Lewis and Thornhill (2003, p291) state:

> **The validity and reliability of the data you collect and the response rate you achieve depend, to a large extent, on the design of your questions, the structure of your questionnaire, and the rigour of your pilot testing.**

The questionnaire was piloted and minor changes made to terminology and structure.

Sample Size

Although the number of completed questionnaires was relatively low (n=28) the respondents were from a highly stratified group covering a number of years of student cohorts. Each respondent had an expert view about the questions.

The size of the sample used was small (n=28) and so the issue of extending the size of this sample would have resulted in the need for a more structured analysis employing the use of computer analysis packages such as SPSS or NUDIST.

The author's acknowledge that the size of the sample restricts the assertions that can be made on the basis of questionnaire responses and have ensured that care has been taken not to make claims on insufficient data (Bell, 1999).

With the small sample of size in this study there is a danger that a high percentage of common views have been translated into conclusions.

Survey Findings

Students were asked to identify what modes of study they considered acceptable, most preferred or least preferred. Table 1 shows the number of times a particular option was selected. The data highlights that evenings only and afternoon and evenings are an acceptable mode of study to the cohort. Online/distance learning is also shown to be an acceptable mode. However, the most preferred mode of study is evenings only. Students were asked to indicate their least preferred mode of learning, results here show that these were intensive course, online distance and in company customised courses. The opportunity was given for students to suggest a combination of the learning modes listed. Acceptable would be a combination of intensive evenings and/or intensive afternoon and evenings. The most preferred were intensive evenings and/or intensive afternoon and evening, intensive online or intensive evenings. The least preferred option was Online/in Company. It should be noted that the response rate to these combinations was low and little significance can be attached to these, however, they do provide a useful understanding in terms of future programme development and delivery.

Table 1 - Modes of study (n=28)	**Acceptable**	**Most preferred**	**Least preferred**
a) Evenings only	12	15	1
b) Afternoon & evening	12	6	6
c) Intensive course over 5 weekends (10 hours per weekend or similar)	9	7	10
d) On line/distance learning	11	2	12
e) In company customised course	7	1	13
Combination of the above	b & c (x 1)	a & c (x 1)	d & e (x 1)
	a & c (x 1)	b & c (x 1)	
		c & d (x 1)	
		a & d (x 2)	

Table 2 presents data ranking the preferred modes of learning specifically relating to the case study, the lower the score the greater the learning preference. Students were asked to rank on a scale of 1-11 their most preferred mode of learning to their least preferred, where 1 is most preferred and 11 is least. The data shows that group analysis, lectures and seminar exercises followed by group discussion are the preferred top three. The data supports Cuthbert's (2005) assertion that the combined use of comprehension and operational/experiential learning is important in helping students to do well in case study work.

Table 2 - Modes of learning, ranking (n=28)	**A&D case study**
Group analysis of case studies	1
Lectures	2
Seminar exercises followed by group discussions	3
Working on past exam papers	4
Question and answer session in seminars	5
Individual analysis of case studies	6
Private study (journal, textbooks)	7
Assignment based assessment	8
Internet research	9
Watching a video/DVD	10
Computer based learning	11

Students were then asked to rank their general preferred mode of learning for the other postgraduate diploma modules, see Table 3. They were asked to rank on a scale of 1-11 their most preferred mode of learning and their least preferred, where 1 is most preferred and 11 is least. The results show that lectures, working on past exam papers, and question and answer sessions are the preferred top three.

Table 3 - Modes of learning, ranking (n=28)	**Other Diploma modules**
Lectures	1
Working on past exam papers	2

Question and answer session in seminars	3
Seminar exercises followed by group discussions	4
Group analysis of case studies	5
Assignment based assessment	6
Private study (journal, textbooks)	7
Individual analysis of case studies	8
Watching a video/DVD	9
Computer based learning	10
Internet research	11

The quality of the teaching was perceived to be generally good by respondents, see Table 4. In particular the Analysis and Decision Case Study and the Planning and Control modules were highly ranked. It should be noted that not all students follow the same pattern of study with some having exemptions from certain modules. The figures shown are actual student responses.

Table 4 - Postgraduate teaching assessment (quality)	**Very good**	**Good**	**Satisfactory**	**Poor**	**Very poor**
A&D Case Study	6	12	5		
International Marketing	2	2	3	3	
Planning and Control	5	12	7		1
Marketing Communications	5	9	3		1

Table 5 presents data relating to the hours of study undertaken outside of the classroom. The results show that 86% of students are studying up to 10 hours on top of their full time work commitments and potentially 6 hours of classroom contact time.

Table 5 - External pressures (hours of study/week) (n=28)	**Students**
1-5	14
6-10	10
11-15	3
16-20	0
>20	1

Table 6 indicates that 43% of students felt that they did not maintain a balance between personal life, work or professional studies. However, 57% felt that a balance was achieved.

Table 6 - Did you maintain balance between personal/work/studies? (n=28)	**Yes**	**No**
	16	12

From Table 7 it can be seen that 93% of students enjoyed the Analysis & Decision group work, likewise 82% felt that they personally benefited from group work. It is interesting to note the difference between enjoying and benefiting.

Table 7 - Analysis & Decision case study work (n=28)	**Yes**	**No**	**Sometimes**
Did you enjoy group work?	26	1	1
Do you feel you personally	23	4	1

benefited from working in a group?			

Table 8 shows that 54% of students felt that small group discussion assisted most with their learning and that 21% benefited from large group discussion whilst 25% preferred individual preparation.

Table 8 - Which one assisted most with your learning? (n=28)	**Individual preparation**	**Small group discussion**	**Large group discussion**
	7	15	6

Table 9 shows that only 50% of students felt that the Postgraduate Diploma in Marketing had made a difference to their career, whilst 43% felt that it had not made a difference. Two students were still completing their studies.

Table 9 - Has the Postgraduate Diploma made difference to your career? (n=28)	**Yes**	**No**	**Yet to complete**
	14	12	2

Discussion

Only three of the cohort had completed learning style questionnaires whilst at SBS. Perhaps this could be considered a weakness in the system but one has to consider just how practical it is to know individual learning styles, particularly at a professional level. Then perhaps it is sufficient to consider Shapiro's 4Ps (preparation, presence, promptness and participation) and observe how students contribute in these areas whilst in the classroom.

In terms of programme development it is interesting to note that online learning is perceived to be acceptable; however there is also an opposition to this as it is the least preferred method of study. This was quite different to a survey carried out amongst CIM branch members, many of whom would have studied for their marketing qualifications many years previously.

The students did not rank in-company delivery very highly and perhaps this is due to a combination of their desire to study outside of their place of work and break away from the inherent interruptions that this introduces, together with pressure and time constraints. This last point is particularly pertinent when considered alongside the 43% who felt they did not maintain a balance in their career, personal life and professional studies. However 57% felt they did maintain a balance and perhaps this reflects cultural changes re the desire for greater leisure time. Most CIM students tend to be young, educated professionals – a generic group which tends to perceive a greater degree of work-life imbalance (Tausig and Fenwick, 2001).

It is perhaps not surprising that students ranked interactive modes of study highly with regard to the Analysis and Decision Case Study. From experience the authors have observed a need for peer support during exam preparation and reassuring stroking from respective tutors. This is supported by Pearson's comment, *"Students that forego plenary sessions and group discussions are likely to fail, no matter how individually clever at analysis they might be."*

Observation from teaching the case study over a five year period suggests that students who do not like the participative, student-centred approach tend to become isolated and have a high likelihood of failing the exam.

The authors' experience reinforces the importance of group work and its importance on the quality of learning. The data presented above supports this with high emphasis placed on the enjoyment and benefit of group work. In addition to this the work of Leenders, Erskine and Leenders emphasises the speed of learning over time in relation to individual preparation, small group or large group discussion and the experiential nature of in class interaction.

With regard to other postgraduate modules, that greater emphasis is placed on lecture input and then interaction in terms of past exam papers, question and answer and seminar work. Here we see a more traditional framework of input then discussion e.g. the technical marketing tools followed by application in the safe environment of the classroom.

Results of a previous survey

A survey carried out by the authors in 2003 and updated in 2009 amongst 1,800 members of the UK Chartered Institute of Marketing Wessex Group, and 28 local directors of SMEs, with 253 responses, found that of those engaged in CPD, 16% attended seminars or exhibitions for learning and 13% were engaged in self-study and reading.

In terms of learning mode preferences for CPD, 16% preferred classroom teaching, 26% on-line self-paced learning and 25% intensive weekends. Distance learning was cited as a preferred learning style by 29% of respondents.

This survey was conducted amongst branch members rather than current students so it might be that the responses reflected preferences developed post-qualification over a long period of time. Furthermore, the CIM case study has the potential to enable students to develop skills necessary for self-directed learning (Kreber, 2001) and this may be why the survey (of members who have already passed the exam as part of their final Diploma) and engaged in CPD activity reveals a high level of interest in on-line, self-paced and distance learning.

Conclusions, Limitations and Implications

The authors acknowledge that the research is exploratory and that greater meaning could be drawn on data from a larger population. The findings suggest that the teaching method employed on the CIM case study is a valid one, meeting students' desires for classroom-base learning which is interactive and benefiting from the value of group work.

It also emphasises the importance of Kolb's learning cycle. Case studies are a valuable method of learning, especially for mature professional students, but lack of participation in group work can lead to failure. The case learning style does not suit students who learning style tends to be predisposed towards structured teaching. However, case studies are idea platforms for relating theory to real-world situations.

Students tend to prefer evening classes to other times and tend to prefer class-based activity to on-line or distance learning, intensive courses or in-company programmes. Results of a previous (2003) study of CIM branch members revealed that distance learning was preferable to on-line self-paced tuition, classroom teaching or intensive weekends, but this was in the context of CPD rather than study for the Diploma. This would tend to suggest recognition of the work-life balance problem experienced by many professional people.

Over 90% of the students surveyed for this current study enjoyed the case study approach used by the authors and 84% preferred small group work to large-group discussion. Tutor support is still very important.

Well over half (57%) of students felt that they were achieving a balance between work, personal life, work and studies. Since the CIM classes at SBS are run in the evenings and this is a preferred time for many of the respondents, it reinforces Tausig and Fenwick's research about the important of schedule control.

Recommendations

Further analysis needs undertaking of the link between success in case study learning and individual learning styles, which itself is a function of students' individual psychological characteristics. The importance of the composition of a study group in terms of different individual team member learning styles also requires further investigation.

Bibliography

Bell, J. (1999) ***Doing your Research Project.*** Third Edition, Open University Press.

Blaxter, L., Hughes. and Tight. M. (1996) ***How to Research.*** Open University Press.

Borg, W. R. (1963) ***Educational Research: An Introduction.*** Longman, London.

Cohen, L. and Manion, L. (1995) ***Research Methods in Education.*** Fourth Edition, Routledge.

Cuthbert, P.F. (2005) ***The student learning process: learning styles or learning approaches?*** Journal of Teaching in Higher Education, 10 (2) 225-249.

Cranton, P.A. (1998) ***No One Way.*** Wall & Emerson.

Dollar, D. (2001) ***Practical Approaches to using Learning styles in Higher Education.*** Community College Review. 28 (4) 28

Easterby-Smith, M., Thorpe, R. & Lowe, A. (2002) ***Management Research.*** Sage.

Finn, J. D. and Rock, D. A. (1997) ***Academic Success among Students at Risk for School Failure.*** Journal of Applied Psychology. 82 (2) 221-234.

Francis, J. (1976) ***Supervision and Examination of Higher Degree Students.*** Bulletin of the University of London, 31, 3-6.

Gill, J. and Johnson, P. (1997) ***Research Methods for Managers.*** Second Edition. Paul Chapman Publishing.

Greenberg, G. and Baron, R. (1995) ***Behaviour in Organisations.*** Fifth Edition. Prentice Hall International Press.

Gross Davis, B. (1993) ***Tools for Teaching.*** Jossey-Bass, San Francisco, CA.

Kolb, D. A., Rubin, I. M. and McIntyre, J. M. (1979) ***Organisational Psychology: An Experiential Approach.*** Prentice Hall Inc. Englewood Cliffs, NJ.

Kolb, D. A. (1979) *Experiential Learning Cycle.* Retrieved 21st September 2005, from http://www.infed.org/biblio/b-explrn.htm

Kreber, C. (2001) ***Learning Experientially through Case Studies? A Conceptual Approach.*** Teaching in Higher Education, 6 (2) 217-228.

Leenders, L. A. M, Erskine, J. A. and Leenders, M. R. (1997) ***Learning with Cases.*** Ivery School of Business.

McKeachie, W.J. (1986) ***Teaching Tips: a guidebook for the beginning college teacher.*** Eighth Edition. D.C. Heath & Company, Lexington.

Pearson, D. (1995) ***CIM Workbook, Diploma, Strategic Marketing Management: Analysis and Decision.*** Second Edition. BPP Publishing Limited.

Phillips, E. M. and Pugh, D. S. (1994) ***How to get a PhD: A Handbook for Students and their Supervisors***. Second Edition, Buckingham, Open University Press.

Saunders, M., Lewis, P. and Thornhill, A (2003) ***Research Methods for Business Students.*** Third Edition, Prentice Hall

Shapiro, B. P. (1984) ***Hints for Case study Teaching.*** HBS Case Services, Harvard Business School, Boston, Mas.

Sharp, J. and Howard, K. (1996) ***The Management of a Student Research Project.*** Second Edition. Gower Publishing Limited.

Tausig, M. & Fenwick, R. (2001). ***Unbinding Time: Alternate work Schedules and Work-Life Balance***, Journal of Family and Economic Issues. 22 (2) 101-119.

Terenzini, P. T. (1999) ***Research and Practice in Undergraduate Education: And Never the Twain Shall Meet?*** Higher Education. 38 (1) 33-48.

Tinto, V. (1997) ***Classrooms as Communities: Exploring the Educational Character of Student Persistence.*** Journal of Higher Education. 68 (6) 599-623.

Wanous, J.P., Reichers, A.E. and Malik, S.D. (1983) ***Organisational Socialisation and Group Development: Towards an Integrative Perspective.*** Academy of Management Review, 9 (4) 670-683.

Watson, W. E., Kumar, K., and Michaelson, K. K. (1993) ***Cultural Diversities Impact on Interaction Process and Performance: Comparing Homogenous and Diverse Task Groups.*** Academy of Management Journal. 36 (3) 590-602.

Since the mid 1990s, schools of business have increasingly offered distance (on-line) learning courses. There are clear signs that this trend will continue, as colleges and universities seek to generate additional revenue by attracting a broader pool of students and, in the current economy, look to reduce overhead by decreasing facilities costs. Convenience is a key factor in attracting students, who, unfettered by the traditional constraints of time and place, can fit on-line courses more easily into their packed personal and professional lives. Faculty members, for their part, benefit from not having to contend with traffic, parking, and inclement weather or to vie with students' electronic gadgets for their attention. Still, both students and faculty are accustomed to the classroom learning platform and old habits die hard. On-line learning *is* different, and students and faculty need to be mindful of the adjustments it entails. In particular, to achieve effective distance learning, it is necessary to bridge gaps in:

1. *time and space* – by creating a culture of community that overcomes students' sense of isolation through the encouragement of faculty-student and student-student communication,
2. *learning style* – by understanding all of the ways our students learn and teaching to as many of these as possible,
3. *personality traits* – by appreciating that individuals with certain traits (*e.g.*, self-discipline, -motivation, and –regulation) adjust more easily to a DL environment and that not all our students have these traits,
4. *self-awareness* – by acknowledging our own biases as to what constitutes effective teaching and learning, and

5. *DL experience* – by recognizing that students who are taking a DL course for the first time may not know what to expect and thus require explicit instructions and frequent reminders of their role in a community of on-line learners.

Management educators whose early on-line experiences were fraught with frustration caused by an inferior instructional-technology infrastructure that thwarted their teaching and their students' learning might have abandoned DL. Instructors who have not yet taught on line may fret that it will be difficult to catch up with peers who have on-line teaching experience; on the contrary, latecomers to on-line learning can reap the usual advantages available to late adopters of any innovation. In particular, they can benefit from improved technical-support infrastructure on their campuses for DL and they can learn vicariously from the mistakes of DL pioneers.

From the students' perspective, there could be issues as to what types of students are suitable for such courses. This research focuses on the personality traits of students with respect to online learning and how it relates to their performance. Online learning implies self-discipline in terms of studying on a person's own time, completing assignments, being able to communicate with others and of course, being able to feel virtually being in a classroom when there is no physical meeting place as such. It is hypothesized that some students would inherently fare better in such an environment and some would not. Using data from online classes at the undergraduate BBA program at an AACSB accredited business school, we examine the relevant personality traits and track these with course performance to develop a normative scale regarding capabilities and traits that good students of online courses should have. The results add to the growing literature in the area of distance learning and provide a better understanding of student traits and also show instructors of online classes which personality traits "work" and which do not in an online course environment.

TRAVEL LEARNING COMMUNITIES AND COMMUNITY BASED SERVICE LEARNING: CATALYSTS FOR INTEGRATIVE LEARNING

Kellyann Berube Kowalski, Department of Management & Marketing, Charlton College of Business, University of Massachusetts Dartmouth, 285 Old Westport Road, North Dartmouth, MA 02747, (508) 999-8327, kkowalski@umassd.edu

Jim Lee, Department of Business Administration, Stonehill College, 320 Washington Street, Easton, MA 02357, (508) 565-1478, jlee@stonehill.edu

Jennifer Ann Swanson, Department of Business Administration, Stonehill College, 320 Washington Street, Easton, MA 02357, (508) 565-1349, jswanson@stonehill.edu

WORKSHOP FORMAT

This workshop will present innovative and nontraditional teaching methods which foster creative and integrative learning. The workshop presenters will give an overview of their experiences with learning communities and incorporating short-term travel and community based service learning projects in management and international business courses. Included in this discussion will be course design, challenges faced with these types of teaching methods, and the learning outcomes that were achieved. Workshop attendees will be invited to share their innovative teaching strategies as well. Finally, attendees will participate in a workshop activity where they will pair up and brainstorm options how to integrate courses and incorporate short-term travel and/or service learning in their courses.

PROPOSAL: "GREENING" A NATIONAL PARK: DESIGNING A SHUTTLE SYSTEM FOR GLACIER NATIONAL PARK – AN INTEGRATIVE MODELING EXERCISE[i]

Beate Klingenberg, Marist College, 3399 North Road, Poughkeepsie, NY 12601, +1-845-575-3000, beate.klingenberg@marist.edu

Zu-Hsu Lee, Marist College, 3399 North Road, Poughkeepsie, NY 12601, +1-845-575-3000, zuhsu.lee@marist.edu

ABSTRACT

This paper proposes the development of an exercise for decision science courses that requires students to integrate several methods to build a comprehensive model for a shuttle bus system for a National Park. Such an exercise can address the need for more comprehensive assessment tools, but also for modeling topics that capture students' interest and creativity. The existing shuttle bus system of glacier National park is used as the base for the model development.

INTRODUCTION

Classes covering basic quantitative decision-making models, such as linear programming, simulation, or queuing models, are often dreaded by students, in particular if offered as a required course of an undergraduate program. The mathematical nature and extensive use of spreadsheet programs is often a fear factor for more qualitatively oriented individuals. Excellent textbooks in the field provide instructors with a multitude of exercises and sample problems on various levels, which are created by modeling real-world business scenarios. While these problems serve the important purpose of showcasing the application of the learned models in the business world, undergraduate students with little or no work experience often have difficulties realizing the importance of these models. Also, most such textbook exercises focus on the practice of just one model, and very rarely offer an exercise that requires the integrated use of more than one method.

The lack of such integrative exercises surfaced at our organization also as a problem for assessment. As a business school that is accredited by The Association to Advance the Collegiate Schools of Business (AACSB), program objective assessment is required. In accordance to the procedures, rubrics were developed and implemented to perform the program objective assessment. Once these rubrics were in place, we experienced that certain objectives, such as quantitative or technology skills, require more than standard textbook problems to be adequately evaluated according to what we formulated as our educational outcomes.

This paper develops an exercise to design a shuttle bus system for a national park, requiring students to practice such methods as linear optimization, simulation and queuing. The intent is to fill the need for integrative assessment tool, as well as to provide a scenario that might be easier for students to relate to. Preferably run as a team exercise, it might also allow the students to reduce their spread-sheet phobia as they can develop the required models jointly with others. Although quantitative course might not always include general class discussions, the proposed exercise also lends itself to an exploration of sustainable tourism.

THE SCENARIO

Glacier National Park is located in the Rocky Mountains in the state of Montana, having received its name from a number of glaciers in the area. It was established in 1910 as the tenth U.S. National Park through a bill signed by President Taft [1]. One of the major attractions accessible not only to hikers but the general public is "Going to the Sun Road", a narrow, winding road through spectacular mountain landscapes that takes private drivers up to Logan Pass. To ease traffic, parking problems and environmental impact of thousands of individual vehicles, the park implemented a shuttle bus system along this road in 2007. Visitors can travel all the way up to Logan Pass, or exit at multiple stops along the way to connect to hiking trails and vistas. Due to the size of the road, and driven by the wish to use environmentally sustainable vehicles, the buses in use are limited and size and can carry only up to 12 passengers at a given time.

THE PROBLEM

Two basic questions need to be answered when implanting the Glacier Shuttle: how many buses are needed and what should the schedule be at a given day. The objective is to minimize the number of buses that travel, while offering transportation to all visitors that want to take it. Many variables influence the answer: the travel time, number of bus stops, expected number of visitors and the probabilities of visitors entering and exiting at the main stop in the valley and the pass, as well as the other stops on the way. The exercise will provide students with a step-wise approach to building the model – starting with a basic scheduling problem and adding other parameter to the model. The attempted end result is a spreadsheet that integrates simulation of visitors, provides information about queuing (service level) and answers the questions on bus quantity and schedule, depending on expected visitors. The exercise should be given after the before mentioned methods have been covered, and several class periods are required to provide guidance for the step-wise model development.

REFERENCES

[1] http://www.nps.gov/glac/index.htm

---

[i] Glacier National Park is running a visitor shuttle bus system since 2007. A visit at the park by one of the authors and perceived inefficiencies are the inspiration for this exercise.

THE LESSONS LEARNED FROM INTERNET BEER DISTRIBUTION SIMULATION GAME

Ki-Young Jeong, Engineering Management, University of Houston-Clear Lake,
Houston, TX 77058, (281) 283-3862, jeongk@uhcl.edu
Alex Monchak, Engineering Management, University of Houston-Clear Lake,
Houston, TX 77058, (281) 480-0156, MonchakA@uhcl.edu
Aung Moe, Engineering Management, University of Houston-Clear Lake,
Houston, TX 77058, (281) 482-2970, aungmoemy@gmail.com

ABSTRACT

This article is to share the educational experience that augments the traditional playing of the Beer Distribution Simulation Game (BDSG). The purpose of this article is to reinforce students' learning on supply chain management through the student-played Internet BDSG. Many previous researchers identified that both demand pattern and communication are important factors in the supply chain management through empirical studies and the computer-played BDSGs, and this became the prevailing conceptual knowledge. However, instead of just accepting this conceptual knowledge, we attempted to evaluate its significance and how students could effectively experience this knowledge. After multiple scenario-based games, the ANOVA test showed that those two factors were not statistically significant. However, we could observe that they affected the supply chain performance results to some extent. The performance was also affected by the learning effect and the strategies that students used. The results and discussions provide useful insights to educators and researchers who use the BDSG.

INTRODUCTION

The Beer Distribution Simulation Game (BDSG) is an excellent "illustrative tool" to teach the knowledge and practice of the supply chain management (SCM). Students and researchers could experience, evaluate, and visualize the conceptual knowledge – e.g. bullwhip effect – learned from their SCM textbook by using diverse versions of this simulation game. For this reason, this game has been extensively used in the academic environment for teaching and research purposes. There have been two groups of research found in the literature.

First group refers to the development of diverse BDSGs and their use for SCM education. Machuca Kaminsky et al. [1] developed a stand-alone game to illustrate the bullwhip effect, the centralization effect, and the lead time effect on the supply chain. Jacobs [2] developed the Internet-version of BDSG, which allowed multiple players to play together. It also allowed the two different demand patterns and a player could graphically observe the bullwhip effect – the increasing variability in orders upstream in a supply chain – during the course of the game. Sparling [3] proposed a teaching plan using the BDSG with which he suggested using the Excel spreadsheet to compute the order and shipment quantities during the simulation game. Siddiqui et al. [4] developed a scenario-based BDSG with the descriptive text guide, allowing the scenario change during the course of the game.

The second group uses the BDSGs for further SCM analysis and research. Kimbrough et al. [5] used the genetic algorithm-based artificial intelligence agent in the beer game, and showed that the agent outperformed the human players and generated better solutions than the best practice solutions known in many cases. Hieber et al. [6] developed the BDSG simulation software and analyzed the impact of the seven different order policies on SCM. Paik et al. [7] used the BDSG computer simulation to statistically analyze and identify the causes of the bullwhip effect. In the second group, most of the authors developed their own computer-played versions of the game to consistently analyze the performance of the supply chain.

Although many previous papers discussed the BDSG, their main focus was on either introduction of the simulation n game or the bullwhip effect – e.g. how to measure and visualize the bullwhip effect, and what causes the bullwhip effect, and etc. However, our main focus in this paper is on the demand and communication – information sharing – among human players, not the computer players. Many previous BDSG articles claimed that these two were important factors affecting the total supply chain performance, and this became the prevailing conceptual knowledge in SCM. However, many of the studies used the computer-played BDSG rather than the human-played game for analysis purposes. Hence students still feel the gap between the conceptual knowledge and the actual practice in SCM. Therefore, in this study, instead of just accepting this conceptual knowledge, we attempt to evaluate the significance of this knowledge in the student-played beer game. The student-played beer game is different from the computer-played game since the first has much more complexities and dynamics in the decision processes than the second, and it is much closer to the real world decision process. Hence, the results from this study could provide many insights to the educators and researchers who use the BDSG. We also believe that students could have a better learning opportunity by filling the gap between the conceptual knowledge and practice.

OVERVIEW OF BEER SIMULATION GAME

In the Internet BDSG developed by Jacobs [2], the supply chain consists of four echelons: factory, distributor, wholesaler, and retailer. It also has two different customer demand patterns: the deterministic and the random demand. A customer orders cases of beer from the retailer; and the retailer orders them from the wholesaler; the wholesaler ships cases of beer to the retailer and orders them from the distributor; and the distributor ships them to the wholesaler and orders them from the factory. The factory brews the beer, and obtains raw materials outside. Each echelon is linked and the beer movement cannot skip the adjacent echelon. Figure 1 displays this supply chain where the flow of order is represented by the dotted line while that of beer is denoted by the solid line. It takes three periods to ship an order of beer from an upstream echelon to a downstream echelon, and the order processing time is one period. There are two important decision variables at each echelon: the quantity of shipment and the quantity of replenishment order.

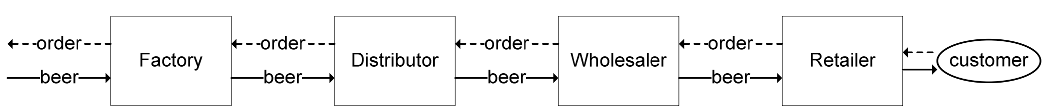


Figure 1. Beer distribution supply chain

The inventory cost per period is $1.00 per case and the back order cost per period is $2 per case. The objective of the game is to meet customer demand for cases of beer through a multi-stage supply chain with cost expenditure on back orders and inventory. The detailed explanation on how to play the game is described in Jacobs [2]

EXPERIMENTAL DESIGN IN TEACHING ENVIRONMENT

The game was played in the graduate level Service and Operations Management (SOM) course in the Engineering Management program at University of Houston at Clear Lake. The twenty students were divided to five teams and each team has four members. Many students in this program have diverse engineering backgrounds, but very few have backgrounds in SCM.

Design of Experiment and Game Preparation

To evaluate the effect of the demand pattern and communication that is sometimes called as the level of information sharing, the following design of experiment (DOE) was set:

- Demand pattern = (deterministic, random)
- Communication = (scenario 1, scenario 2, scenario 3) where
    - Scenario 1: Each player could not share any information
    - Scenario 2: The retailer could tell its customer demand to other team members
    - Scenario 3: all information could be shared and discussed.

- In scenario 1, each player in a team makes an independent decision without any communication with other team members. In scenario 2, the partial information sharing is allowed – the customer demand to the retailer is provided to other team members at each period. Scenario 3 allows the complete information sharing within a team and each member freely discusses the problem with other members within *one and half minutes*.
- In each scenario, students could use any ordering strategies they want as in the real world business. The strategies may change over time depending on their judgment.
- These experiment factors could be considered with *two independent simulation games* – game 1 with the deterministic customer demand and game 2 with the random customer demand. For each game, three scenarios are repeated.

RESULTS AND ANALYSIS

Initially we practiced the game for five periods, and then completed the three scenarios for twenty five periods, each with the deterministic demand. After one week, we repeated the same scenarios with the random demand. During the week time interval, the students were asked to review the results of the first game, and prepare for the random demand game.

Table 1 summarizes the class-wide results of the total supply chain cost for. The last two columns computed the average cost and standard deviation per team. According to this, two teams – teams 3 and 5 – show better performance than other teams in terms of cost minimization and consistency. Each team's total cost per scenario is plotted in Figure 2 where the x-axis represents the sequence of scenarios played, and the y-axis shows the total cost for that scenario. In the x-axis, ds1, ds2, and ds3 represent the scenarios 1, 2 and 3, respectively, for the deterministic demand, and rs1, rs2, and rs3 represent the scenarios 1, 2 and 3, respectively, for the random demand. As seen in Figure 2, the performance among teams significantly varies. Initially we thought that the game with the random demand could be more difficult than the game with the deterministic demand. However, teams 1, 3, and 5 generated better results with the random demand than with the deterministic demand. Overall, each team's performance got stabilized as time progressed, and the results from the second game with random demand were more consistent than the first game. Hence, we believe that the learning effect positively contributed to the second game. The chart in Figure 2 also showed the consistent trend between the cost and communication level – the more communication level the teams have, the less total cost they have.

Team	Deterministic demand			Random demand			Mean	Standard deviation
	Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	Scenario 3		
1	109498	11417	5164	12061	7508	5887	25256	42367
2	6384	2389	4105	15738	3604	2021	5707	5151
3	7820	6791	4079	5231	2776	2357	4842	2184
4	13398	2019	1275	4229	4600	1972	20371	15847
5	109498	11417	5164	12061	7508	5887	4528	4520

Table 1. Total supply chain cost

The results in Table 1 were analyzed with ANOVA in Table 2 where both the demand pattern and the communication level were statistically analyzed at the significant level (α) 0.05 using the statistical software MINITAB. Surprisingly, all factors including interactions were not statistically significant. Although they were not significant, the communication had the largest impact on the total supply chain cost (e.g. whose *p-value* is 0.178).

Source	*df*	*SS*	*MS*	*F-test*	*p-val*
Demand	1	2.24104E+07	2.2410435+07	0.05	0.822
Communication	2	1.60359E+09	8.01793560+08	1.86	0.178
Interaction	2	4.36540E+08	2.18269958+08	0.51	0.609
Error	24	1.03573E+10	4.31552228+08		
Total	29	1.24198E+10			

Table 2: ANOVA table for beer distribution simulation game

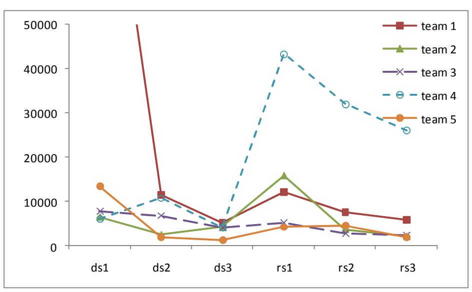


Figure 2. Cost per time

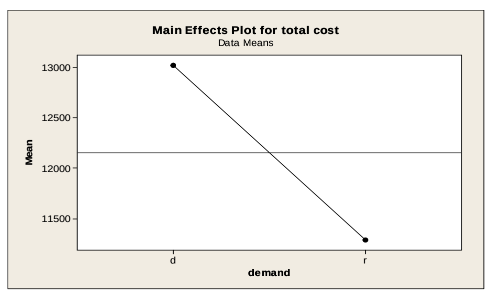


Figure 3. Demand effect

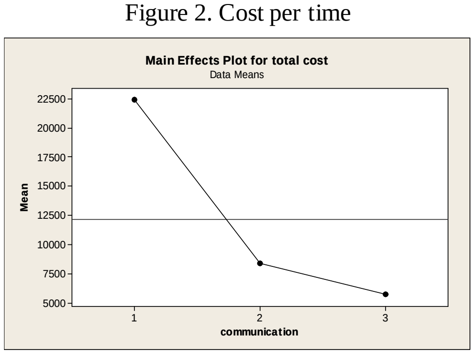


Figure 4. Communication effect

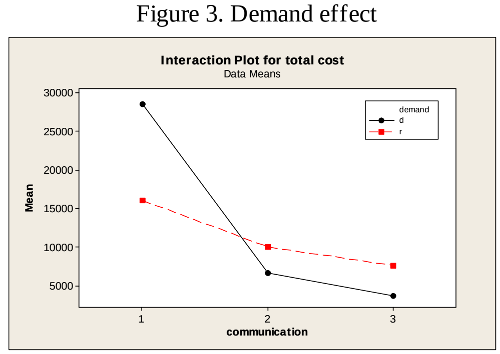


Figure 5. Interaction effect

Figure 3 displays the chart showing the effect of the demand pattern – total cost reduced when the random demand (denoted by r) was used when compared to the deterministic demand (denoted by d). The cost also reduced when more communication was allowed – see Figure 4 for communication effect. There was minor interaction effect between demand and communication as shown in Figure 5. Although none of these was statistically significant in our study, we could still observe that the demand and the communication level to some extent affected the supply chain performance in this game.

TEAM LEVEL ANALYSIS

The team level analysis could also be interesting since each team may have different communication effectiveness and use different strategies during the game. We evaluated each team's performance by using two-way ANOVA at the significance level (α), 0.05 and found out that no demand and communication was significant except the team four's demand that was identified as significant at *p-value* 0.036.

In Figure 6 teams 1, 3 and 5 performed better at the deterministic demand while teams 2 and 4 performed better at the random demand. While the effect of demand was mixed among teams, the effect of the communication level was consistent at all teams as seen in Figure 7. Most teams performed better when more information was given although the impact was not statistically significant.

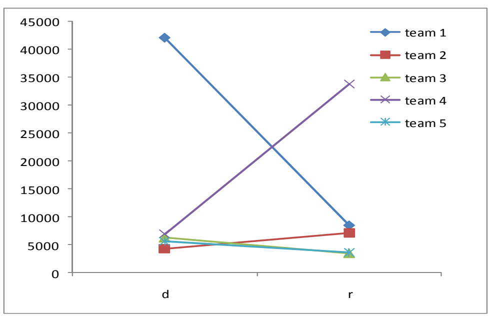


Figure 6. Demand effect at each team

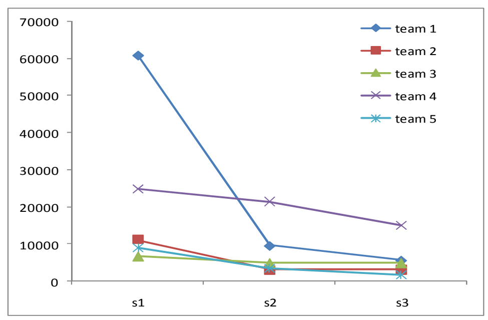


Figure 7. Communication effect at each team

During the course of the game, all teams were allowed to use any strategy they wanted. In fact, many teams attempted to make a balance between supply and demand at each echelon based on their own strategies. Some used a dynamic strategy where they changed their ordering strategies over time while others used a constant strategy throughout the game. Both teams 3 and 5 attempted to use the chase strategy throughout the game. Within this strategy each echelon attempted to synchronize its current shipment to the order it received from the previous period. This may explain why those two teams had relatively lower average cost and smaller standard deviation as observed in Table 1. That is, in addition to the learning effect, we also believe that the strategy implicitly affected each team's performance and the teams that used the consistent chase strategy showed the better performance.

DISCUSSIONS

In our experiment with students, we initially thought both demand pattern and communication level could be significantly important factors affecting the performance of the game. However, several statistical analyses showed that none of these was statistically significant in this study. Nevertheless we could still observe a trend - when more information is available, the performance gets better. We observed that the conceptual knowledge was supported in the interaction effect chart in Figure 5.

We learned that the performance was affected by the learning effect as the game was repeated. After twenty five periods, we have computed the mean and the standard deviation of the customer order for each game. For the deterministic game, the mean (μ_d) and the standard deviation (S_d) were (μ_d, S_d) = (7.2, 1.63), which generated the coefficient of variation (C_d), 0.23 defined by S_d/μ_d. For the game with the random customer demand, the mean (μ_r) and standard deviation (S_r) were (μ_r, S_r) = (7.2, 2.86) and its coefficient of variation (C_r) was 0.43. That is, although the random demand case has much more variation than the deterministic demand case, this randomness was not sufficient enough to overcome the learning effect according to the main effect analysis in Figure 3. This indicates that it would be useful if the game allows the user- controlled the demand variation.

We also noticed that the results of the game could vary depending on the strategies used. For example, students may use a constant or a dynamic order strategy per period. We recognized that the two teams with the consistent chase strategy had better results than others.

CONCLUSION

We have used the Internet-based BDSG to see whether the conceptual knowledge – demand pattern and communication affect the performance of the supply chain – could be realized and evaluated in the student-played beer game. The BDSG developed by Jacobs [2] was used for this purpose where students play as a decision-maker at each echelon. We used two levels of demand pattern and three levels of communication for this experiment. The results showed many variations from team to team. However, none of those factors was recorded statistically significant in this study. Nevertheless, the trend in the main communication and the interaction effect charts supported the conceptual knowledge to some extent.

During the course of the game, it was noticed that the performance was also affected by the learning effect. The learning effect could be evidence of the effectiveness of this game as a teaching tool. We also identified that the consistent use of the chase strategy improved the performance throughout the game regardless of the demand pattern. Overall, although it was not statistically significant, we observed to some extent that the conceptual SCM knowledge worked through the experiments.

In summary, we experienced that the BDSG by Jacobs [2] is a good learning and practice tool for SCM. Students are humbled by the experience of participating on losing teams in the six scenarios, but learning the supply chain knowledge through this scenario-based simulation game. In the future it would be desirable to provide the user-controlled customer demand variations and the hybrid game mode where students play with computers using a specific order strategy for more effective learning.

REFERENCES

References available upon request from the authors

CAREERS IN INFORMATION TECHNOLOGY: CAN SOCIAL NETWORKING MAKE A DIFFERENCE?

Esther E. Klein
Department of Management and Information Technology
St. Francis College, 180 Remsen Street, Brooklyn Heights, NY 11201, (718) 489-5233, eklein9@aol.com

Corinne Smolizza
Department of Management and Information Technology
St. Francis College, 180 Remsen Street, Brooklyn Heights, NY 11201, (718) 489-5451, csmolizza@optonline.net

WORKSHOP FORMAT

Presentation and discussion on careers in information technology with a focus on social networking sites and the role they play on attracting and retaining women in IT careers.

SUMMARY

"Social network sites (SNS) are a recent computer-mediated communication (CMC) technology receiving attention in the popular press for their ability to enroll new members at an astonishing rate" [5, p. 1]. MySpace launched in 2003 using the slogan "a place for friends," was acquired by Fox Interactive Media in 2004. MySpace membership has been growing exponentially – at 80 million members at the end of 2006 [1]. Facebook, with membership initially limited to Harvard University students, was launched in 2004. In 2006, Facebook was opened to the public, currently reporting membership of 175 million worldwide users, making it the number one social networking site [6]. LinkedIn was designed and positioned to attract the young professionals. The site launched in 2003 with the five founders inviting 350 of their most important contacts to join. At the end of the first month in operation, LinkedIn had a total of 4,500 members in the network. Today, LinkedIn has more than 48 million users spanning more than 200 countries [2]. MySpace, Facebook, and LinkedIn are among the most popularly used social networking sites that we will examine. Other sites such as Flickr, Friendster and Twitter do attract a small percentage of users whose influence we will also explore.

Most social networking research has focused on student use [4]. Rapleaf [3] reports that women and men use social networking sites equally but differently. Women use social networking sites such as Facebook and MySpace to build and nurture relationships, whereas men use LinkedIn and are less likely to spend as much time nurturing relationships as they are "transacting." Despite increased presence of older demographics on social networks, the Rapleaf study reveals that younger people still continue to dominate social media space. As such, can social networking sites be used to attract and retain women in IT careers?

REFERENCES

[1] "A place for friends" A history of MySpace (2008, August 14). Retrieved November 11, 2009, from Random History website: http://www.randomhistory.com/2008/08/14_myspace.html
[2] LinkedIn (2009). Company history. Retrieved November 11, 2009, from LinkedIn website: http://press.linkedin.com/history#2009
[3] *Rapleaf* (2008. June 10). *Study of social network users vs. age.* Retrieved November 5, 2009, from Rapleaf website: http://www.rapleaf.com/company_press_2008_06_18.html
[4] Skeels, M., & J. Grudin (2009). When social networks cross boundaries: A case study of workplace use of Facebook and LinkedIn. *Group '09,* May, 95-103.
[5] Sledgianowski, D., & S. Kulviwat (2009). Using social network sites: The effects of playfulness, critical mass and trust in a hedonic context. *Journal of Computer Systems,* Summer, 74-83.
[6] Stross, R. (2009, March 8). When everyone's a friend, is anything private? *The New York Times*. Retrieved November 11, 2009, from The New York Times website: http://www.nytimes.com/2009/03/08/business/08digi.html

A DECISION SUPPORT SYSTEM FRAMEWORK for VENDOR MANAGED INVENTORY in SUPPLY CHAIN MANAGEMENT

Eric Y. Cheng, SUNY-Canton, (315) 379-3904, cheng@canton.edu
Gordon Yen, Hong Kong Polytechnic University, (852) 2766-7370

ABSTRACT

This paper presents an ongoing research on the initial development of a decision support system (DSS) framework for Vendor Managed Inventory (VMI) in general consumer goods supply chain. A conceptual intelligent DSS framework provides a holistic framework to perform analytical assessments of the data generated from VMI software. First, a fundamental objective hierarchy for VMI and the associated objectives and influencers were developed. Second, the inventory decision module which determines the appropriate stock levels and the sales forecasting module which provides periodical forecasts at each product level were proposed. Third, functions include parameter estimation, updating, and forecasting system performance assessment were presented. Lastly, by integrating the modules and functions described above, the DSS development framework was proposed. This paper contributes to the retailing industry by proposing a more efficient and effective way of improving their supply chain's efficiency in order to gain competitive advantage.

Keywords: Decision support, supply chain management, vendor managed inventory

INTRODUCTION

With the continuing weakness in the global economies has lead to the reduction in profit margins for many industries. Moreover, the strategic importance of the supply chain management is increasingly recognized because companies have faced a growing cost competition and customer demands. As a result, more companies are collaborating with their supply chain partners to improve their supply chain's efficiency to gain competitive advantage. One of the solutions that has been widely adopted is vendor managed inventory (VMI), also known as vendor-managed replenishment or consignment inventory. However, not all retailers are able to successfully implement VMI. According to Schenck and McInerney[5], "One of the key issues for failure in the implementation of VMI is the lack of integration with key decision-support and reporting systems within retailer and manufacturer organizations" It's often that the data is printed out and then re-keyed into order-management systems. The valuable data that can help influence a trading partner's production, replenishment and shelf strategies continues to remain in the VMI software, underutilized [7]. This practice continues to hinder the growth of VMI; thus critical mass, and the real benefits, are not being realized by most of the industry." This illustrates the importance of the integration of the decision support system (DSS) with the VMI process.

This paper presents a DSS framework for VMI between general consumer goods retailers and their manufacturing suppliers. To develop a general VMI DSS framework, existing works on VMI in the apparel supply chain and the characteristics of the general consumer goods supply chain were examined. The framework was developed based on knowledge from literature review and observations of the current conditions in industries.

LITERATURE REVIEW

Vendor Managed Inventory

According to Tyan[6], the motivation behind a VMI system is that both the retailer and supplier work together to maximize the competitiveness of the supply chain and the most obvious benefits of VMI are inventory cost reduction for the retailer and total cost reduction for the supplier. Moreover, the productivity and service level improvement can result in larger profit margin and increase in sales. Achabal et al. suggested that a key business motivation for developing VMI replenishment systems is to develop a deeper partnership between the vendor and the key retail accounts. Some of the specific goals and benefits proposed by Achabal et al. [1] for this system are to:

Goals:

- Give the retailers' customers the best opportunity to purchase the vendor's products.
- Help the retailers manage their inventory more effectively.
- Assist the vendor in production scheduling.

Benefits for the retailer:

- More effective inventory management and less uncertainty regarding inventory turnover and customer service levels. The VMI system provides a way to set and achieve performance targets for both these goals.
- A cost-effective way to obtain sales forecasting and inventory management services. As the vendor's analysts implemented the system across many retailers, economies of scale were achieved in both the development and the customization of the models. This lead to a VMI forecasting system that was more accurate and developed at a lower cost than could be realized by any individual retailer.

Benefits for the vendor:

- VMI provides a method for the vendor to increase the availability of their brand in stores, relative to competitors' brands, and still meet the retailers' budgetary open-to-buy constraints.
- Retailers' orders are often misleading data for production planning. Orders do not provide accurate information about which merchandise sells more rapidly and which styles stocks out in midseason, for example. Furthermore, less popular styles and colors are typically sold eventually through markdowns. Relying on actual sales data also prevents the "bullwhip effect" [4], that occurs when time lags, coupled with batch orders from the retailer, tend to amplify demand fluctuations as they go up the supply chain.
- VMI also reduces the opportunity and incentives for gaming, for example, retailers sometimes intentionally inflate orders when product supplies are limited and proportionally allocated by the vendor.

DSS FRAMEWORK for VMI in RETAILING

As the retail supply chain involves many smaller units, such as each retail outlet of a chain store, that operate much like an independent small business, the stages proposed by Chaudhry et al. [2] can be used as a reference for developing DSS for VMI. In order to design a DSS for VMI, it is important to fully understand the objectives and influencers for VMI in General consumer goods retailing.

Objectives and Influencers of VMI

The ultimate objective of VMI is obviously to improve profitability. Since VMI is a collaborative process involving the retailers and their suppliers, it is therefore necessary to consider the profitability of both groups of participants.

Profitability of Retailer

Profitability of the retailer is determined by sales and cost of the retailer. In the context of VMI, the profit of the retailer is mainly affected by the ability and effectiveness of the inventory management system to optimize inventory cost and maximize sales. To maximize sales, it is necessary to maintain customer service level in terms of availability of goods, maintain at least sufficient quantity of goods to achieve proper presentation at the store shelves, and launch seasonal promotional programs to stimulate and trigger sales around seasonal events such as holidays. Different consumer products may require different strategy but these three factors are universal for most of the general consumer products.

The ultimate question that a VMI DSS should address is "What quantity of each SKU should be ordered for each ordering period?" This is closely related to the current stock level and the target stock level. Whereas the current stock level is a known figure, it is necessary to consider various other factors in order to determine the target stock level. These factors are 1) open-to-buy budget of the retailer; 2) inventory turnover requirement of the retailer; 3) target customer service level at the retail outlets; 4) minimum inventory level required to properly present the goods at the retail outlets; 5) available production capacity of suppliers; and the forecasted demand until the next order period. To come up with a demand forecast, it is necessary to consider factors including 1) baseline sales or de-promoted sales which is the sales figure without special promotion effort; 2) seasonal effect that captures the periodic variations in sales that are affected by seasonal events such as holidays and back-to-school; 3) marketing efforts include all factors affecting sales that are controllable by the retailer, such as price, advertising and in-store presentation.

Based on the above and using the method suggested by Clemen & Reilly [3], we can develop a fundamental-objective hierarchy for VMI in general consumer goods retailing as shown in Figure 1 and an influence diagram as shown in Figure 2. Examination of the objectives and influencers described in Figure 1 and Figure 2 shows that they are applicable to the apparel supply chain as well as most general consumer goods.

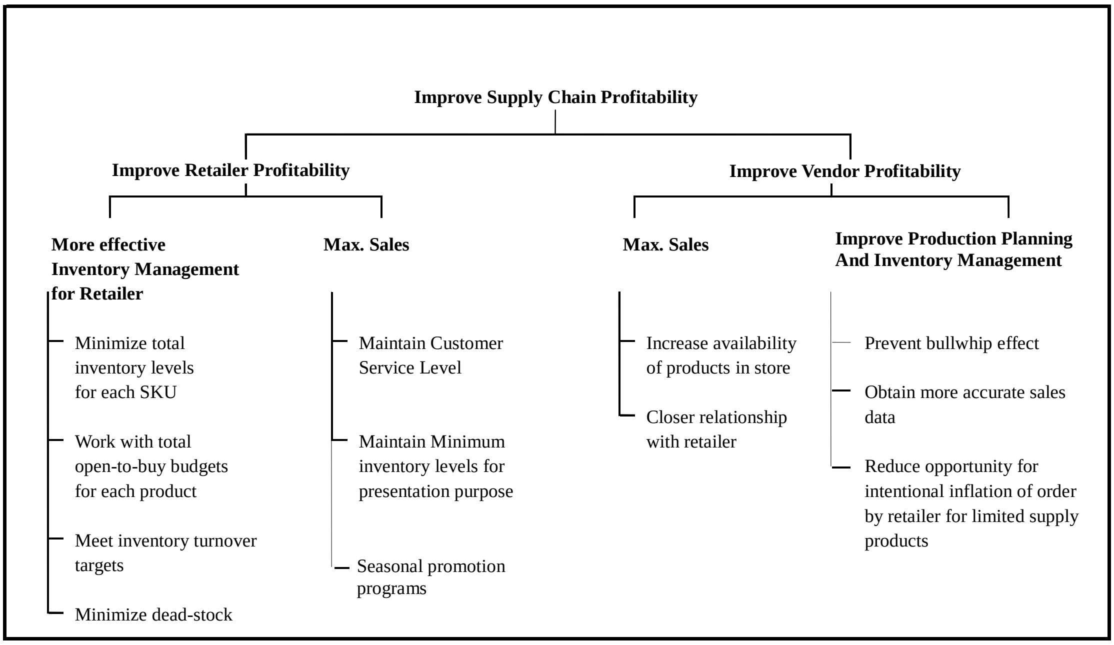


Figure 4, Fundamental-objective hierarchy for VMI:

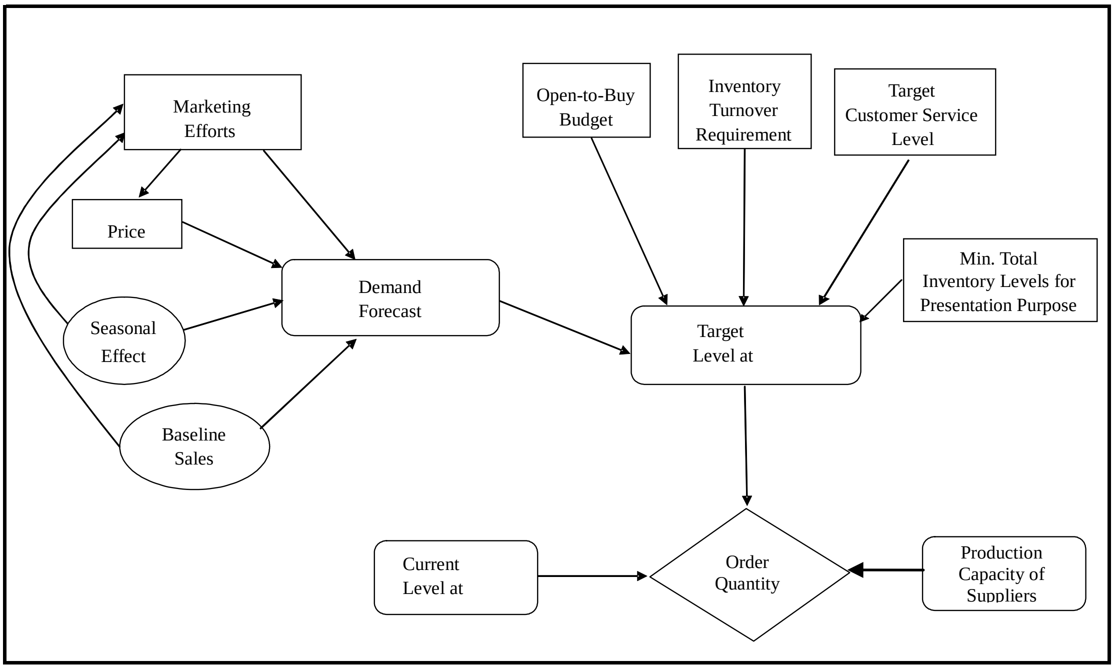


Figure 5, Influence Diagram for General Consumer Goods VMI

Parameter Estimation and Updating

Parameter estimation provides initial values for the parameters of the forecasting model. After developing a forecasting model, linear regression can be performed using historical sales data that includes all seasonal variations. Stepwise regression should then be applied to identify the coefficient values that had significant sales impact. The frequency of sales forecast should be selected based on the operations of the retailers, particularly how the sales data is captured and updated.

Parameter updating smoothes the initial values for the parameters of the forecasting model generated by the parameter estimation process to adjust for changes in sales patterns over time. Parameters estimated by regression on the past sales data would not be able to reflect any changes in market conditions or promotional strategy for the current season. Also in the case of new products, the regression estimates can only be based on sales of similar, but not identical, products in the previous periods. As such, various methods have been used for parameter updating in retail sales forecasting systems. Some examples are exponential smoothing and discounted least squires. Testing should be done using historical data to determine the effectiveness of these methods for a particular situation.

DISCUSSION

Whereas VMI and DSS are both well researched topics, there appears to be no specific framework developed for the development and implementation of consumer goods VMI DSS. If the result of this research is able to derive a universal framework for such purpose, it would be a significant academic contribution with real commercial value. However, it would be impossible to include all possible categories of consumer goods in a single study. Therefore, the result of this study can only be considered applicable for the most common consumer goods which supply chains are same or similar in nature to apparel, consumer electronics, toys, and furniture. Moreover, different supply chains or even different companies in the same industry may have very different influencers for inventory decision making that could affect the compatibility of the framework. It may be difficult to generalize in this regard. Nevertheless, we believed that this framework can bring the world one step closer to a more effective and efficient way of developing and implementing VMI DSS for consumer goods retail supply chain.

REFERENCES

[1] Achabal, D. D., McIntyre, S. H., Smith, S. A., & Kalyanam, K. (2000). A decision support system for vendor managed inventory. *Journal of Retailing, 76*(4), 430-454.

[2] Chaudhry, S.S., Salchenberger, L., Beheshtian, M., (1996). A small business inventory DSS: Design development, and implementation issues, *Computers Ops Res.* Vol. 23, No. 1, 63-72.

[3] Clemen, Robert T., and Reilly T., *Making Hard Decisions*, Duzbury. ISBN: 0-534-36597-3, 2nd Edition, 2001.

[4] Lee, Hau L, V. Padmanabhan and Seungjin Wang (1997). "Information Distortion in a Supply Chain: The Buillwip Effect," *Marketing Science*, 43 (4): 546-558.

[5] Schenck, J., McInerney, J. (1998). Applying vendor-managed inventory to the apparel industry. *Automatic I.D. News* 14 (6), 36-38.

[6] Tyan, J., & Wee, H.-M. (2003). Vendor managed inventory: a survey of the Taiwanese grocery industry. *Journal of Purchasing and Supply Management, 9*(1), 11-18.

[7] Yao, Y., Evers, P.T., Dresner, M.E. (2007), "Supply chain integration in vendor-managed inventory", *Decision Support Systems*, Vol. 43 No.2, pp.663-74.

AUTOMATED NEGOTIATION SYSTEM ARCHITECTURE

Mu-kun Cao, School of Management, Xiamen University, Fujian, China 361005, +011866-13625019295, caomukun@yahoo.cn
Robert Chi, Information Systems Department, California State University, 1250 Bellflower Blvd, Long Beach, CA 90840, (562) 985-4238, rchi@csulb.edu
Ying Liu, Information Systems Department, California State University, 1250 Bellflower Blvd, Long Beach, CA 90840, (562) 985-4554, yliu2@csulb.edu

ABSTRACT

Electronic commerce has changed the way businesses interacting with consumers and peers. For both B2B and B2C transactions, it becomes more and more important to make the traditional negotiation price mechanism automated and intelligent [1]. Automated negotiation has become the core of the next generation intelligent e-commerce. The main principle is using software agent technology to make the negotiation process partially or fully automated. The purpose is to improve the efficiency of online negotiations and reduce costs of transactions. In recent years, the research on automated negotiation system has been given high priority by researchers around the globe [2].
Keywords: Automated negotiation system, multi-agent system, service-oriented negotiation, web services.

INTRODUCTION

Traditional research in automated negotiation is focused on theory about negotiation protocol and strategy. However, the application of automated negotiation system has lagged far behind. Negotiation protocol defines the interaction rules between agents. Negotiation strategy defines the sequence of actions taken by agents in the negotiation process based on decision-making models. The research results of protocol and strategy are very rich and represent the major research effort in automated negotiation.

At present, one severe problem for the study of automated negotiation is how to convert the theoretical results into practical applications. Although there are many research achievements about protocols and strategies in the field of automated negotiation nowadays, realization and real application of automated negotiation system still has a long way to go [3, 4]. So far, there is scarcely any automated negotiation system that can be applied in e-commerce [5], which has been a bottleneck for the research of automated negotiation. Many theories, models and algorithms cannot be verified without a practical application platform, thereby constraining the further development of automated negotiation research. In fact, this kind of situation is widespread in the entire e-commerce oriented automated trading applications [6].

PROBLEM STATEMENT

While some systems have been widely cited in literatures, such as the Kasbah, Tete-a-Tete, ADEPT, AuctionBot, eMediator, MAGNET, MATE, etc., however, strictly speaking, they are not automated negotiation system because software agent technology in these systems is primarily used for automated trading functions, such as product selection, price comparison [7] and so on. Even some applications involved in negotiation are mainly for auction. This situation can be attributed to the following reasons:

First, current research lacks negotiation systems. Scholars pay more attentions to the economic meaning of negotiation in the research of automated negotiation. They are more interested in building a variety of mathematical models to describe the negotiation and designing decision-making process models. These studies are built on the basis of an assumption that there has been a software platform, agent, to support the operation of these mathematical models. However, what on earth is the negotiating agent? How does it follow the negotiation protocol and implement the negotiation strategy? These are critical questions toward the practical application of automated negotiation system. Unfortunately, these questions have not been paid sufficient attentions.

Second, current research needs development methods for automated negotiation system. Automated negotiation system (ANS) is essentially a multi-agent system; therefore its development is inseparable from the software agent technology. Given that software agent technology and multi-agent system development method are not mature, the development of automated negotiation system has no technology roadmap to follow.

Finally, negotiation system research needs a feasible application mode. Most negotiation systems are developed in academic environments such as laboratories in universities. They are usually experimental systems used to verify a particular model or theory and face many technical and security problems. There is a big gap between the experimental systems and the practical commercial applications. At the same time, colleges and universities find it difficult to make the experimental systems meet practical application requirements. Consequently, these systems only stay in the experimental stage. Moreover, many such systems were abandoned later, which is a great waste. Therefore, we need to introduce a new application mode to support the whole process from research to application.

SOFTWARE AGENT TECHNOLOGY FOR AUTOMATED NEGOTIATION SYSTEM

This paper discusses a technology roadmap for the development of automated negotiation system using the software agent technology and proposes a practical application architecture using SOA and web services technology for the automated negotiation system.

A service is a software component that can be accessed via a network to provide functionality to a service requester. The term service-oriented architecture refers to a style of building reliable distributed systems that deliver functionality as *services*, with the additional emphasis on loose coupling between interacting services. Technically, the term SOA refers to the *design* of a system, not to its implementation. We regard SOA as an *architectural style* that emphasizes implementation of components as modular *services* that can be discovered and used by clients.

The service-oriented architecture is very useful for the application of automated negotiation system. As mentioned above, there are many difficulties, such as security and credit, to make the automated negotiation system to be applicable. In fact, in most cases, the enterprises and individuals do not want a software system, but just a negotiation service. On the other hand, it is more convenient for the developers to just provide a service to the users rather than to deploy software system in the user site, for there will be less work of maintenance, operation and so on. Therefore, we can make the automated negotiation system as a service and deploy it using an enterprise application integration framework based on service oriented architecture. To do so, the ANS can take advantage of the infrastructure such as network, security, transaction and so on provided by the whole integration system and benefit from the existing customer resource.

Negotiation Service generally has the following characteristics. 1) It is individually useful, or it can be integrated and composed to provide higher-level services. Among other benefits, this promotes re-use of existing functionality. 2) It communicates with their clients by exchanging messages: they are defined by the messages they can accept and the responses they can give. 3) It can participate in a workflow, where the order in which messages are sent and received affects the outcome of the operations performed by a negotiation. 4) It is completely self-contained, or depends on the availability of other services, or on the existence of a resource such as a database. In the simplest case, a negotiation might be perform without needing to refer to any external resource, or it may have pre-loaded all the data that it needs for its lifetime.

Moreover, the automated negotiation system base on SOA will have the following good cross-platform features: 1) Flexibility: A negotiation service can be located on any server, and relocated as necessary. As long as it maintains its registry entry, prospective clients will be able to find it. 2) Scalability: negotiation services can be added and removed as demand varies. 3) Replacing ability: provided that the original interfaces are preserved, a new or updated implementation of a service can be introduced, and outdated implementations can be retired, without disruption to users. 4) Fault tolerance: If a server, a software component, or a network segment fails, or the negotiation service becomes unavailable for any other reason, clients can query the registry for alternate services that offer the required functionality, and continue to operate without interruption.

REFERENCES

[1] C. Bartolini, C. Preist and N. R. Jennings, A software framework for automated negotiation. *Software Engineering for Multi-Agent Systems III: Research Issues and Practical Applications*, Springer Verlag 2005, LNCS 3390: 213-235,.

[2] A. R. Lomuscio, M. Wooldridge and N. R. Jennings, “A classification scheme for negotiation in electronic commerce”, Int Journal of Group Decision and Negotiation 2003, 12 (1): 31-56.

[3] Resinas M, Fernandez P, Corchuelo R.. A conceptual framework for automated negotiation systems. Proceedings of Intelligent Data Engineering And Automated Learning - Ideal 2006, 4224:1250-1258

[4] He MH, Jennings NR, Leung HF. On agent-mediated electronic commerce. IEEE Transactions on Knowledge and Data Engineering. 2003, 15(4): 985-1003

[5] Resinas M, Fernandez P, Corchuelo R. (2006). A conceptual framework for automated negotiation systems. Proceedings of Intelligent Data Engineering And Automated Learning - Ideal 2006, 4224:1250-1258

[6] Debenham J. (2005).Foundations for automated trading - It's the information that matters. Proceedings Of Database And Expert Systems Applications, 2005, 3588: 534-543

[7] He MH, Jennings NR, Leung HF (2003). On agent-mediated electronic commerce. IEEE Transactions On Knowledge And Data Engineering. 2003, 15(4): 985-1003

ENHANCING CORPORATE ENVIRONMENTAL MANAGEMENT PERFORMANCE WITH GREEN INFORMATION TECHNOLOGY AND VIRTUALIZATION

Edward T. Chen, U. of Massachusetts Lowell, (978) 934-2756, edward_chen@uml.edu

ABSTRACT

The image of the company from the perspective of customers and the public in general must be carefully monitored as it relates to the environment and the use of natural resources and energy. The demonstration of effective strategies that allow for "greener" and more ecological awareness, will gain the respect of customers, businesses, stockholders, and other concerned groups. The protection of our environment is also a major agenda of firms today. This paper draws insights from the literature and discusses how "Green IT" along with the concepts of "Virtualization" can provide for better organizational, operational, and environmental outcomes.

Keywords: green IT, virtualization, environmental information system, energy saving

INTRODUCTION

The large number of personal computers that are present in businesses and the high quantity of servers that are available twenty-four hours and seven days use tremendous amounts of energy and space. Strategic and tactical solutions can be put in place that can effectively reduce costs for organizations and allow for "greener" overall results [5][18]. Additionally, by using various Virtualization techniques for hardware, software, and systems, many daily operations can be made more efficient. Incorporation of various strategies can lead to new ways of conducting business such that higher cost savings and less environmental impact can be realized [9][17][21].

Organizations must always attempt to resolve financial, logistical and environmental issues in order to remain competitive and to be perceived superior in the industry and in the public eye. In order to be viable as an organization in this century, concerns for the environment must be considered [24]. Failures to recognize or act upon these looming challenges will ultimately destroy companies and economies [25][30]. Creative tactics must be taken into account to strengthen the corporation as well as the world around it [1][3][6]. While organizations are exploring creative alternatives to save billions of dollars, they need to make our environment healthier and more stable [10][13][19][23][33].

GREEN INFORMATION TECHNOLOGY

There has been much discussion over the past several years about our environment and the effects of global warming. Organizations such as Greenpeace and influential people such as Nobel-prize winner Al Gore and his documentary, "An Inconvenient Truth", have certainly put much focus on these topics [2]. Perhaps it is this awareness that has made each of us more conscious of the potential outcomes that may ensue if we continue to ignore our environment. This knowledge has moved from our own consciousness to the organizations in which we work. Green Information Technology (IT) is the term used to capture this awareness and to study the potential ways to solve the problems with pollution and consumption of vast amounts of energy in the computing world [11][27].

The concerns for economic viability, social responsibility and the effect on the environment are all encapsulated in this concept of Green IT. The outcome of developing solutions that can be more friendly to the environment and at the same time more economical for the corporation has led many to believe that the ultimate goals of an organization are simply to take advantage of the concept by exploiting and touting the concept of being “green” for marketing purposes and for the potential profit gains that could be realized. Although this argument may contain some truth, one should ask the following question. Is it not better that companies be concerned for the environment, whether it be for legitimate reasons or selfish reasons, so that the ultimate goals of society can be realized? A corporation has an obligation to both its shareholders and to the economic viability of society in general [29]. If the outcome of being “green” allows a company to be more profitable either by means of better efficiency or by better public relations, then so be it, as long as the ultimate worldly consequences are reduced [6][10].

VIRTUALIZATION

The term “Virtualization” has multiple meanings depending on the context of the discussion, but it can be summarized as the creation of logical pools of IT resources not linked to physical devices [26]. When used in the context of servers, virtualization refers to the use of special software which allows systems administrators to install multiple operating systems on the same physical piece of hardware. When used in the context of desktops, it refers to the hosting of individual user sessions or desktop operating systems at a central location. When used in the context of storage, virtualization refers to the dissemination of data over a wide range of physical locations in an attempt to achieve optimized access across a large geographical area.

Virtualization, in a broad sense, is the simulation of a real physical entity. In the information technology field, this term is used to describe a means of hiding the physical characteristics and technical details of a system from the user [9][17]. For example, virtual servers are computer systems that can provide multiple users with a perception that they are being used exclusively, but in fact, the same computer system may be serving the needs of several users. The system has the power to serve the needs of many applications but physically consists of one set of hardware. This allows a better utilization of equipment such that idle time of a computer is minimized. The result of this Virtualization is better utilization of hardware and thereby an ultimate reduction in the use of energy to run the systems [9]. Costs are reduced not only by the reduction in necessary hardware and infrastructure but also in costs to power the systems and the surrounding infrastructure [15].

HELPING THE ENVIRONMENT

We need to find ways to lower our power consumption and reduce the amounts of heat and other pollutants into the atmosphere. Additionally, the amount of hardware that is used and is then discarded, due to newer technology or better performance, must be reused or recycled. It is estimated that carbon dioxide emissions from servers and other computing components account for 0.75% of all annual amounts of emissions. Contributing to this amount, other IT and telecommunications equipment constitutes at least two percent to this output [12]. Additionally, alone the United States uses twenty-five percent of the entire world's electricity of which one and one-half to three percent is absorbed by data centers alone. This amount of energy has doubled in the past five years and is expected to more than double again in the next five year period.

In order to combat these frightening numbers, Green IT initiatives must be developed and deployed. Utilization of more efficient cooling systems and the use of better insulation techniques will aid towards the reduction of data center power usage and production of excess waste [14]. Applications of Virtualization to reduce the quantity of servers in data centers thereby reducing the amount of energy consumption for power and cooling of servers will also be better for the environment [22][34].

Corporation Benefits

Going Green is not only good for the environment and the world in general but is also good for business. Employing the various concepts of Green IT can strengthen a corporation and provide side benefits that are not obvious upon immediate evaluation. In the past decade, the emphasis has been on buying more and more computer power with higher and higher processing capabilities. It was assumed that the necessary infrastructure, including power, cooling, and building space was readily available and affordable. This is not the case today. Infrastructure is the predominant concern and limiting factor for most businesses today. As businesses attempt to grow they are challenged by the excessive costs of energy, lack of space, government regulations, and pressure from society [14][36].

As businesses have grown, their reliance on computer systems and data has grown as well. Google, for example has close to a half million servers in its network. Their continued expansion is requiring more powerful and faster computing power in addition to more storage capacity for their data [34]. This ultimately results in higher energy costs and more facilities to house the equipment. Unfortunately, many of the existing facilities that exist today are not equipped to handle the massive amounts of power required or the cooling capabilities necessary to keep the facilities at appropriate temperatures. The heat generated by all of this equipment must be removed properly in order to insure proper reliability. It is estimated that reliability of hardware decreases by fifty percent for every eighteen degrees above seventy-eight degrees Fahrenheit [34]. These levels of poor reliability would be devastating to organizations, especially those that rely on data warehousing as a primary operation for their business. The reality of the situation is that it is now costing a substantial amount of operating budgets to power and cool equipment. These costs are overshadowing the costs of actual equipment. It is interesting to note that the problems with facilities and the ability to power and cool equipment is at critical stages for almost 50% of businesses in operation today [14]. These limitations cause businesses to have limited growth potential even though the need for their services and products is in high demand. In today's competitive and global environment, issues like this are a thorn in a firm's operations.

In addition to the operational obstacles that hinder business growth, governments are raising concern for the environment and are putting pressure on IT organizations to identify and develop solutions to the power consumption and environmental issues that their businesses are contributing to. One way or another, corporate officers, customers or governments will eventually demand that all businesses deal with the issues that are confronting our society [29]. The cost of doing business will certainly need to anticipate these changes. Establishing a plan for the infrastructure needs to deal with these issues is of upmost importance. Ensuring that this plan meets the rapidly expanding demands of the future is a business practice that will have long lasting benefits [31][33].

Cost Effectiveness

Corporations are learning new ways to cut costs by experimenting with various conditions and measuring the overall benefit. It is important that companies share this information so as to provide others with similar benefits. Many of these cost optimization techniques potentially bring firms to a competitive edge. Corporations may hold those cost reduction practices as confidential and proprietary properties. On the other hand, there is sometimes a contradiction in doing the most admirable thing for a business versus doing the right thing for society. Other times, there is somewhat of an overlap and both entities benefit from the outcome [14][22].

Reduction of costs can be accomplished in a variety of manners. There are quite a few obvious approaches that have been touched upon in preceding sections of this document. Other approaches are a little more creative but equally provide high potential outcomes. A few approaches are listed below for the reader to consider and ponder upon.

It has been determined from a study performed by a consortium of companies that are dedicated to maximize the efficiency of data centers, that at least sixty percent of the power required to cool IT equipment is wasted due to inefficient infrastructure [22]. Costs saving can be enhanced by simply controlling the costs of power and cooling by methods previously described and by making certain that the way in which facilities are cooled and insulated meets appropriate standards. Efforts should be made to improve the facilities layout and thermal infrastructure. For example, taking the hot air produced from equipment and moving it to other areas that require heat to maximize and utilize a by-product of the IT equipment in an effective manner. Moving a facility to another part of the world where alternative energy can be used rather than fossil fuel based energy is both a cost effective tactic as well as an environmentally conscious activity [14][22].

It has also been estimated that most servers and desktop computers are only used eight to fifteen percent of the time. Unfortunately, when computers are idle, their power draw is not. The hardware itself, whether idle or active, consumes sixty to ninety percent of typical workload power requirements [14][32]. A simple way to address these issues and to cut costs is to make certain that systems purchased have a low power mode of operation such that when not on active duty, they consume less energy. Automatically powering down servers during periods of low demand can decrease energy costs by as much as eighty percent [14][32].

Finally, Virtualization is one of the most important and perhaps the most common method of cutting costs by effectively reducing the number of servers required to perform the work in demand. Even though it may be relatively easy to determine the costs of power and cooling for IT servers and equipment, it is not always easy to justify a large scale restructuring in terms of return on investment. Reduction in servers by using Virtualization techniques can save on the cost of utilities. But, there is also a cost associated with replacing the lower performance servers and a cost associated with buying the Virtualization software and supporting infrastructure. These costs are not insignificant and will be scrutinized by management. A solid business case must be developed to demonstrate the benefits both in terms of the bottom line and in terms of the long term strategic viability and benefits [9].

CORPORATE SHIFT TO SERVICE-BASED IT

It is always interesting to look back into the past and to compare the current state of affairs with those of a period a long time ago. In the world that we live in today, it seems unlikely that events of another century would have any applicability to the present, however, this is not the case. Take for example the case of electric power generation at the turn of the twentieth century. For years, each industrial organization developed the necessary power to supply its plant by using natural resources such as waterfalls to generate electricity or steam engines to drive electric generators. In years to follow a shift from self supplied energy to utility supplied grid energy occurred [8].

At first, the economic and logistical benefits were not very obvious. Today most would agree that the business of producing energy should be left to the utility companies [7]. It is far easier to purchase the necessary power than to actually generate it. It seems highly likely that in the years to come, IT infrastructure could shift to follow a similar utility type model. This type of revolutionary change will change the entire IT landscape for corporations. As discussed above, the issues of power consumption, cooling and infrastructure in general could be shifted to a utility company. The net effect is that a business can re-focus on its business and not have to be concerned with the burdens introduced by IT complexities. “IT is shifting from being an asset companies own to a service they purchase.” [4][8].

In most scenarios, the cost to purchase services from utility companies is too economically feasible to ignore. The reality is that, a transition to this model will take time. There is evidence, however, that a shift may lie in the horizon. The current atmosphere suggests that the time is coming when utility based capabilities will be available. Shifting to centralized supply tends to be more attractive when overcapacity is coupled with redundant functionality. This is currently the case with most large scale IT firms. If a system were to be put in place, it seems likely that a shift would occur. Obviously, this shift would take time and effort to fully accomplish [4][8][16].

Once a move to utility based computing occurs, the need for personal computers and other standalone devices will potentially diminish. The concept of providing devices that allow connection to a service based utility for computing will take the place of PC’s. This will allow a centralized infrastructure to be put in place and thereby eliminate the massive number of systems deployed in businesses that are essentially sitting idle on desks the majority of the time. Cost savings, power savings, and environmental savings will be enhanced astronomically [35].

WORTHINESS TO BE GREEN

Many argue that the additional costs of attempting to be "Greener" counteract any potential business benefit that a firm may realize. So, the real question is why a firm would want to invest in being green if it cannot justify it from a cost basis. Research has shown that greener organizations tend to exhibit better economical performance than less green organizations but there is no supporting evidence that it is a result of their level of greenness [20][28].

It may not be cost justified to become green for the sole and selfish reason of the bottom line, but it may be justified based on ethics and a general overall concern for the environment. Providing benefits to society may not be tangible assets for a company to possess but the good-will that is demonstrated will certainly provide priceless satisfaction and credibility to the organization. This demonstration of concern will solicit customer loyalty and will ultimately reward the firm in ways that may not be easily traceable. In summary, it seems quite obvious that ultimately it pays to be green for the sole purpose of demonstrating good-will and showing concern for society. Obviously there are many challenges for a company to overcome to become Green.

The important parts are related to the success of the business and the overall benefit to the environment. Decisions and tactics should be put in place to maximize the benefit of both these objectives. One must also be aware that there is no perfect solution that will satisfy the needs of all. In fact a positive outcome in one area is probably offset by a negative outcome in another area. Establishing business objectives that advance the firm's best interests and allow a firm to play an important part in society should be the ultimate objective. The rest will follow by a course of natural events.

CONCLUSION

Over the next few decades it will be interesting to see the overall ambition of corporations in response to the environmental challenges that our world is being exposed to. This coupled with the fact that costs of doing business are accelerating at rapid rates makes it interesting to forecast the future business landscape. Corporations and individuals bear the challenge and respond with wise decisions regarding the use of vital resources and the production of environmental toxins. We can find a happy medium where profits can be made at little or no expense to the surrounding world. Our technical advancements and desires will make the world a better place for this generation as well as our great grand children's generation. Now is the time to respond to the challenges. As executives and managers in the businesses that we work in, it is our responsibility to force change that can minimize our footprint on the world. It is also our responsibility to facilitate the profitability of the firms in ways that ensure success but do not compromise our surroundings. Both agendas are achievable.

Though there is a lack of a comprehensive research literature for the specific topic of Green IT and Virtualization, this paper draws on research results and insights from a broad array of disciplines. Much of the information of Green IT and Virtualization was provided in the sources found as listed in the references. Several of the topics researched in this paper gave the readers a better comprehension what server Virtualization is all about and how changes in the IT landscape provide substantial benefits to organizations and at the same time to help our challenges on earth.

"References available upon request from author."

Parameters of Green IT

Rajath Chouta
Department of NSSA,
B. Thomas Golisano
College of Computing &
Information Sciences
Rochester Institute of
Technology
Rochester, NY 14624
rsc5726@rit.edu

Tae Oh
Department of NSSA,
B. Thomas Golisano
College of Computing &
Information Sciences
Rochester Institute of
Technology
Rochester, NY 14624
tom.oh@rit.edu

Jungwoo Ryoo
Information Sciences and
Technology
The Pennsylvania State
University-Altoona
Altoona, PA 16601
jryoo@psu.edu

Young B. Choi
Department of MIS & CIS
College of Business
Bloomsburg University of
Pennsylvania
Bloomsburg, Pennsylvania
17815-1401
ychoi@bloomu.edu

ABSTRACT

We currently have a number of standards and solutions to measure and improve the degree of being "green" for a network or IT equipment is. Some of these standards and solutions overlap in their coverage while others address entirely different aspects of green information technologies. Many authors of these standards and solutions also claim that their approaches are superior in helping the IT industry become more environmentally responsible. This paper investigates the similarities/differences among these standards, strengths/weaknesses and coverage, and tries to categorize them. In addition, we attempt to identify standards highly relevant to designing a green computer network infrastructure in general. We believe that it becomes much easier to implement sustainability within the IT sector once all the different parameters associated with a greening process are established in a coherent fashion.

1. Introduction

The green revolution has taken over everything in its path, including the IT sector [10]. Because of the aggressive progress toward greening effort, new green technologies are being developed without a clear direction or consensus on how the corresponding green standards should be developed to be truly meaningful.

This paper provides brief descriptions of all these standards and compares them for their coverage, effectiveness, and relevance to greening IT industry in general and computer networking in particular. Identification of all the major parameters of green IT will help the balanced development of necessary green standards by exposing overlaps and missing standards.

The ever shrinking resource pool as well as ever increasing carbon footprint caused by the IT industry is our main motivation for this research. Large networks like data centers account for a significant portion of carbon emission today, and there is an imperative to handle this situation so that it does not go out of control and cause irreversible damage to the ecology that envelopes us.

The major parameters existing today are discussed at first, which is followed by hardware solutions and software solutions respectively to expedite the process of going green. The brief analysis, including the advantages and disadvantages of each solution are discussed in the respective sections. The areas that are not touched upon by the parameters are discussed at the end, followed by suggestions that can be implemented in the future standards.

2. Energy Efficiency Measuring Parameters

2.1. *Energy Star*

Energy Star happens to be one of the oldest environmental programs devised with co-operation from the government of a country in an effort to reduce the pollution caused by electronic appliances [9]. It attempts to dictate carbon emission limits that different appliances should follow in order to receive the Energy Star certification. All devices that carry this particular label are at least 10% more efficient than appliances certified by the federal standards [4].

Originally, only USA was actively making use of this certification in order to correctly identify and label devices that are considerate to the environment. Today, this is a reputed symbol recognized in countries like Australia, Canada, Japan, Taiwan and many countries across Europe [4]. Energy Star was initially conceived as a parameter to measure power plant efficiency, but it has far outgrown this initial application. Today, one can even find homes that have been Energy Star Certified. Additionally, each component of a large enterprise network has its own kind of Energy Star Certification. Since the parameters are different among different devices, the devices cannot be compared for their Energy Star rating, but can be done so within the same family of the devices.

Energy Star parameters are updated every year, and it is required that the manufacturer should also renew their certification if they want to continue using the Energy Star symbol [4]. Hence, the stringent emission norms coupled with regular updates do make this a versatile standard that is very important in the process of greening IT or any network hardware.

Figure 1. *Energy Star Logo [3]*

2.2. Power Usage Effectiveness (PUE)

This is an important term used by corporations to describe the efficiency of their data center. It is defined as the ratio of the total amount of power used by the data center to the power delivered to the IT equipment [14]. Although in lay man's terms, it sounds like quite a simple definition, the measurement of this parameter is quite challenging. The power delivered to the IT equipment refers to the power consumed by the storage devices, the network equipment, the Keyboard, Video, Mouse (KVM) switches, the monitors, the workstations, and any other device that has been used to monitor or control any of the equipment in the data center. The total power used by the data center is another large value and encompasses aspects like the Uninterrupted Power Supply (UPS), the cooling unit, generators, Power Distribution Units (PDUs), computing and storage nodes, miscellaneous loads like lighting and other accessorial equipment. This parameter has a number of uses although limited by its nature. Two important advantages of calculating the PUE are identifying whether your data center matches up to other similar data centers in power consumption and the possibility of improving the data center efficiency. The difficulties lie in that you need to collect the statistics of all the different devices accurately. Otherwise, this parameter won't be of much use. A perfect data center will have a PUE of 1.0, which is almost impossible. There aren't any data centers currently that can achieve such a high PUE. If your calculations reveal a PUE below 1.0, then there is something wrong with the acquired data or the calculation. One of the lowest PUEs ever been reported is one of Google's data center with an average PUE of 1.19 and an all time best rating of 1.11 [2].

2.3. Corporate Average Data Center Efficiency (CADE)

It is clear that PUE doesn't consider the aspect of utilization and places emphasis purely on the power efficiency parameter. To bring in the factor of utilization as well, one needs to use Corporate Average Data Center Efficiency (CADE). With this parameter, you can clearly identify how many of the resources are actually being used. It is possible for a company to have a high PUE such as 1.2, but the actual usage of resource could be a different picture altogether. When you bring in the utilization parameter, it becomes clear as to whether or not the data center is truly making use of all the resources optimally. This allows the users to determine whether it is possible to turn off certain servers at certain times of the day. The parameter is determined as a product of facility efficiency and IT asset efficiency [6] where facility efficiency is the ratio of energy delivered to IT and the energy drawn from the utilities. IT asset efficiency is the average CPU utilization across all the servers.

The combined formulae according to [5] is given as -

CADE = Facility Efficiency X IT Asset Efficiency

The parameters here are further defined as -

Facility Efficiency = Facility Energy Efficiency X Facility Utilization

Asset Efficiency = IT Energy Efficiency X IT Utilization

A study has revealed that the average CPU utilization in data centers is 6% [6], which clearly shows large amount of resources are being wasted. The centers could use virtualization to save these resources. Hence, having a good PUE rating is seldom an indicator of a good data center. With companies adding more data centers and demanding more energy, it has become vital for efficiency to increase to a higher level. The processing demand for data centers tend to rapidly grow in scale, but the solution is not to add more servers. Instead, companies ought to improve the efficiency of existing infrastructure.

2.4. Data Center Infrastructure Efficiency (DCiE)

The Green Grid consortium [1] coined another term in an effort to measure the efficiency of the data center. It is determined by dividing IT equipment power by total facility power, and expressed as a percentage value. One can observe that DCiE is the reciprocal value of PUE.

$$DCiE = 1 / PUE$$

Hence, it is another form of representation of PUE itself. With these two parameters, it is possible to determine whether or not the data center operations are in fact increasing the efficiency of the data center over time. From the DCiE percentage value, one can also estimate how much power is required to be delivered in order for the data center to work efficiently. If the DCiE is around 50%, that means you need to send twice the amount of power theoretically necessary for the data center to work optimally [1]. The challenge with both these parameters is that in many cases, a data center might be housed as a part of a building. While estimating the power being consumed by all the non-IT equipment not connected to the data center, there will be a margin of error in the calculation. Additionally, it can be tedious for the system administrator to do the calculations on their own as they might not always get the right value when it comes to measuring the power outside the data center. It is not clear what would be the ideal DCiE value as not all data centers have to mandatorily measure this parameter. However, a DCiE of at least 65% is considered being efficient [1].

3. Hardware Oriented Parameters

3.1. *Thin Clients*

The concept of a thin client is quite straightforward. With this device, the core components of a computer are retained, and all the accessorial components are shifted onto a central server. Core components are defined as video, basic I/O, networking, small amounts of RAM, a processor, a monitor and some form of input device [12]. Hence, you can make use of these machines if you want to save power and restrict the number of resources required to sustain the enterprise network. It has been observed that a typical user does not need the processing power offered by a full blown desktop computer. For light activities like Web surfing, checking e-mails and using simple applications, you don't need a full-fledged desktop machine [7]. Using a thin client can conserve huge amounts of power in the long run. In fact, it has been estimated that a thin client uses only 7 watts of power compared to the 25 watts that a regular desktop computer consumes [12]. Thin clients are very easy to acquire, but there are some challenges to successfully implement them. The biggest challenge is that these thin clients need to have good amounts of support and that the users should be familiar with how to use them. In most of the places that the thin client solution has been proposed, the challenge is to get the users to work on these terminals instead of a Windows-based solution [12]. In certain cases, the initial phase of shifting the network towards a thin client model has been time consuming, but after the transition to the thin clients, the network has performed flawlessly [7]. The limited computation power means that only a certain amount of the network can really be shifted, and the whole network cannot be made to run only using a thin client as such.

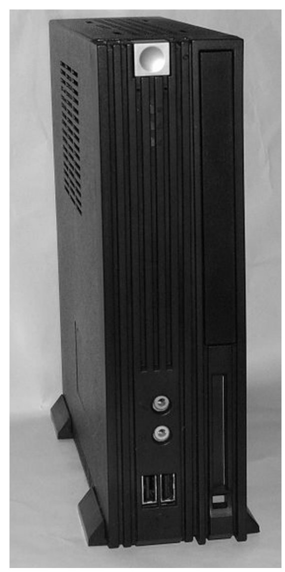


Figure 2. *A Thin Client Machine (Source: Wikipedia)*

3.2. Relocation of Infrastructure

An alternate approach for green computing would be to just relocate the existing infrastructure to a more efficient location. This would not necessarily be a very cost effective solution. However, if upgrading network is considered, relocation of infrastructure is probably a viable solution. Typically, locations for large networks like data centers are selected only after exhaustive research and detailed data has been collected about a particular physical location. If, on the other hand, you are looking at a newer facility to house the network, you can consider some other factors that can help you make the choice. For example in Upstate New York, the cold temperature of a region can help cool down the data center. These are features that you couldn't have possibly thought of previously and perhaps can be put into effect in the newer facility. Also, the heat generated in the data center can be put to better use if the heat is sent to the ventilation system to disperse in the other rooms of the facility, which require heating. Such concepts have been used only recently and are still being actively researched on. Some other new developments being considered while choosing a site for the enterprise network are its accessibility from different locations as well as its proximity to a major form of transportation like air or sea. Another parameter is whether the building allows expansion at a later stage or if the network has to be migrated entirely during the time of expansion.

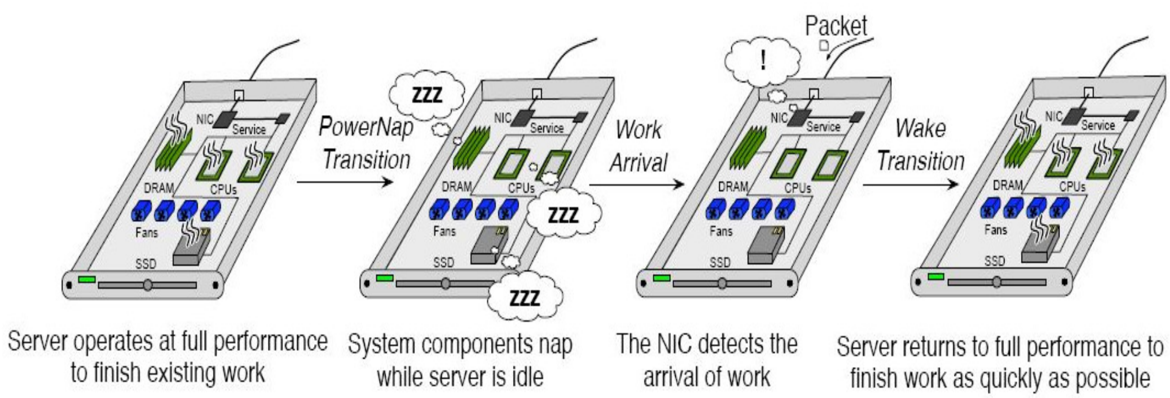


Figure 3. *PowerNap in Action [8]*

3.3. Using Power Smart Devices

This represents a viable alternative that can be immediately adopted as it has advantages for the environment as well as for the organization. With efficient usage of power, the benefit to the environment is pretty straightforward. Another reason for a viable solution is that it even helps the organization save money in the long run. As how Thin Clients use only 28% of the power that a typical desktop requires [12], using similar power-aware devices can significantly reduce the utility bill. Ultimately, the organization is interested in getting this number down over everything else. Therefore, they are willing to ensure that it happens. In this way, the concept of *PowerNap* [8] is quite useful. With this, devices are put in a very low power sleep mode. This mode enables them to take minimal power, and only the NIC is maintained at an active state throughout.

When the NIC detects traffic, it instantly turns on the devices and makes sure that the required components are brought out of their low power mode to the necessary mode to service the request introduced in the network. This is just one of the many different power aware solutions out there. The key is to choose a solution that does not need drastic changes to be incorporated in the network and that will not cost the management too much to implement. Also, replacing all older Power System Units (PSU) with newer and more efficient ones will help improve the efficiency rating of a data center. Typically, PSU ratings of 80 and above are preferred. There are parameters that are being used in other devices as well like cooling, heating and peripheral devices. These parameters can also help determine if your devices are meeting the approved rating. Furthermore, some standard energy ratings like Energy Star are very beneficial and help consumers tremendously by making it easier for them to decide the apparatus to buy. Ensuring that the network is running only on such devices can significantly improve the efficiency rating of the network on the whole.

4. Software Oriented Parameters

4.1. Virtualization

The quickest way to implement a green solution and improve all the energy related ratings of your data center would be to embrace virtualization. Almost every data center in North America today has some form of virtualization in their network [11]. Once you are aware of all the benefits associated with this technology, you will understand why it is so popular. Considering the fact that servers are seldom utilized to an efficient level, it would make sense to use a single computer to host more than one service. Virtualization takes advantage of this and enables a single system to run multiple services and even more than one operating system simultaneously. The cost savings, in terms of associated costs as well as investment costs are impressive. The carbon footprint will immediately drop, which should give the company tax breaks; something that can be used as a selling point to make a company start using virtualization. The cost of implementation is minimal as you are only required to pay for the license of the software. Your existing infrastructure might be powerful enough to handle virtualization without any problems. For companies that are cash strapped but need to expand their resources, implementing virtualization is a logical and cost effective measure. A less need for new hardware means money can be spent on training people to work with this technology. Moreover, the other aspect that encourages people to switch to virtualization is that administration of the network becomes simpler. You don't have as many hardware components to take care of. When the computer is working fine without any problems, then the virtual server will work flawlessly. The only disadvantage that companies might find is that the software licensing cost can sometimes be prohibitive. In addition, virtualization must be implemented in a network that has resources for it. Virtualization on an already burdened network can be catastrophic.

4.2. Remote Management

Remote management is a concept rapidly being implemented in enterprise networks to facilitate control from multiple locations. There are benefits for having such a feature in the network. To a certain degree; it even facilitates the move to implement renewable solutions with the enterprise network. If the data center is located in an area that doesn't have favorable access or is in a location that doesn't have adequate trained professionals to be on site round-the-clock, remote management makes it possible to take control from any compatible locations, remotely reboot the systems or even resolve any software issues that might occur in the network. For companies planning to incorporate virtualization, a remote management solution can help manage these virtual machines and ensure that no single resource is being strained or that the load is being balanced across all the available systems [13]. With load balancing, even the power system units (PSUs) of the machines can work at optimal efficiency. In fact, effective network management is only possible when you are taking care of your network continually, even if you are not around the vicinity of the network. Although in most cases the network will take care of itself [13], periodic checks are necessary to just be sure of how the networks are running and verifying that everything is working like it is supposed to.

Another key benefit with remote management is that costs are significantly reduced, and the money saved in hiring round-the-clock technicians can be used for something else. For the corporations, big or small, implementing any solution is directly related to the costs, and if money is saved, they are more likely to implement the solution. There are a number of providers in this remote management field, which offer total remote control of the network. Unlike the previous days when remote management was only restricted to certain devices, today's modern networks are capable of being completely controlled remotely, including even basic peripheral devices.

4.3. Using Power Saver Modes

Another way to save power without having to invest in anything extra for your network would be enabling the devices on your network to use the power saver modes which could be equipped in devices already. Almost all the components used to build networks today have some form of power conservation protocol. The challenge lies in being able to implement these modes without affecting availability in any way. A solution that has been previously mentioned (PowerNap) can be used in this regard [8], but you can even put into effect power saving profiles with existing resources, which are normally turned off in most devices. Additionally, if you are implementing remote management, it is possible to enable power saving remotely and put devices that are not being used to sleep. In this manner, there can be energy savings while at the same time being environmentally aware. Besides, even the longevity of the devices is more likely as they are not taxed for prolonged periods and can in fact be put to rest now and then.

Two challenges are immediately apparent for implementing this solution. The first challenge is to ensure that the responsibility of doing this task is not given out randomly. It is important to find someone who has the knowledge to analyze the network traffic and predict the traffic patterns. The second challenge would be in identifying types of power saving modes from different devices and utilizing the modes as much as possible. Many devices have hibernation features, but this is seldom made use of owing to the long time needed to recover from the mode. In fact, it negates any of the advantage that you would get by saving power as the cost involved in starting the device and getting it back up at optimal efficiency will be quite high. Once a reliable network traffic pattern is available, the power saver profile can be defined and utilized efficiently.

5. Challenges

5.1. Budgetary Challenges

For any company, the only reason that they will look at change would be to see whether or not the company profits from transition to this alternate option. As a result, if the cost of the transition is not justified, the company will not be in a position to adopt it even if it's advantageous to the environment. Thus, the money factor also needs to be considered in the equation to be sure that the management will not have any problem with the proposed solution.

5.2. Other Challenges

There are also some other challenges that one might have to look into before jumping into this green concept. For example, a known problem with implementing virtualization is that it becomes harder to isolate power consumed by each virtual machine. As the virtual machines are loaded on a single physical machine, the power drain caused by each virtual machine is very difficult to isolate.

Another problem that one might have to look into is that while there are a plethora of standards out there, no single one is sufficient.

Since green practices are different in different parts of the world, it becomes challenging to truly measure whether or not the network is running at acceptable efficiency as compared to another network located elsewhere, in a different part of the world.

Education remains a concern with going green. No matter how beneficial it might be for the environment, users need to adapt to the standard on their own and have the urge to save the environment. If the people working with the equipment fail to use the environment saving measures of the devices, the entire exercise is a futile one even with the implementation of all of these measures.

Since IT equipment manufacturers don't have a particular measure of being green to adhere to, it gets tricky when you are mixing technology from one provider with another one. This problem could be resolved by using as many products as possible from the same manufacturer. Unless there is no other option, using products from two different manufactures with the same functions should not be encouraged.

6. Future Work

So far, we have come across a number of different greening parameters being used by different companies. In the near future, there will be a need to extend these measurement standards and solutions to standalone systems and PC users as well. Today, a bulk of the measurements is based on large scale computing and massive networks. When end users are more educated on how they can also contribute to going green, carbon emissions will be reduced significantly. Just like how Energy Star [4] is used to determine energy efficient devices, it is important to bring out similar standards that can rate the devices on other important parameters.

Some examples of this include cooling efficiency and supporting resources requirement (like what might be needed in order to get a device to work optimally). The key is to maintain harmony between different components and ensure compatibility and optimal efficiency. If there is a global standard to which equipment manufacturers can adhere, it can be much easier to implement green solutions for real world cases.

While green research as of today is mainly limited to technological problems, the next issue to address is to develop regulations on going green and to enforce them. Unless there is a law that is enforced on people, few organizations would make significant changes in the way they do their business.

7. Conclusion

This paper shows that there are a sizeable number of parameters addressing how "green" a computer network or a firm can be. There are also a significant number of companies investing in greening efforts. However, it becomes quite challenging for a large multi-national company to adopt a universal policy since the environmental rules are not fixed in all countries.

In addition, it is of importance to observe that end user knowledge is vital to ensure that all policies are adhered to. Drafting policies that are not adhered to makes no sense. By being frugal with the usage of resources like paper and power, one can contribute significantly to going green. Policies of a company are crucial when it comes to such matters. If there is an overly relaxed atmosphere at the workspace, the policies are not going to be followed as effectively.

REFERENCES

[1] Belady, Christian; *Green Grid Data Center Power Efficiency Metrics: PUE AND DCIE,* obtained from http://www.thegreengrid.org/~/media/WhitePapers/White_Paper_6_-_PUE_and_DCiE_Eff_Metrics_40_December_2008.ashx?lang=en.

[2] *Data Center Efficiency Measurements, http://www.google.com/corporate/green/datacenters/measuring.html*.

[3] *Energy Star,* obtained from http://en.wikipedia.org/wiki/Energy_Star

[4] *ENERGY STAR® — The Power To Protect The Environment Through Energy Efficiency,* obtained from http://www.energystar.gov/ia/partners/downloads/energy_star_report_aug_2004.pdf.

[5] *Fast Guide: Data Center metrics and standards* from http://www.searchdatacenter.com.

[6] Hoover, J., Nicholas; *Measure Data Center Efficiency like Car Fuel Efficiency.*

[7] Kissler, Simon; Hoyt, Owen, *Using Thin Client Technology to Reduce Complexity & Cost,* SIGUCCS '05 Conference, California, USA.

[8] Meisner, D., Gold, T. G., Wenisch, F., T; *PowerNap: Eliminating Server Idle Power;* ASPLOS '09, Washington DC, USA.

[9] Mifflin Mariette, *What is energy star?,* obtained from http://housewares.about.com/od/majorappliances/f/energystar_RO.htm.

[10] Ramon Mata-Toledo and Young B. Choi, Green Computing, 2010 McGraw- Hill Year Book of Science and Technology, 2010.

[11] *Reduce IT cost and go Green,* VMWare solution brief.

[12] Ritschard, R., Mark; *Thin Clients: Make Them Work for You;* SIGUCCS '09 Conference, St. Louis, Missouri, USA.

[13] SRENA, *Introduction to Console Management and SRENA,* Outpost Sentinel White Paper, 2002.

[14] Szalkus, Mark; *What is Power Usage Effectiveness?,* obtained from http://ecmweb.com/design_engineering/power_usage_effectiveness_1208/.

UTILIZING ADVANCED ANALYTICS TO IDENTIFY NETWORK DISRUPTIONS THAT IMPACT PATIENT LENGTH OF STAY METRICS

Abstract

Numerous performance metrics have been established to measure the efficiency of healthcare providers. In order to more fully understand and manage these metrics, analysts must leverage available data resources describing the network of activities that underpin them. Interdependencies between departmental activities exist within the patient treatment process, where inefficiency in one element of the patient care network of activities can adversely affect outcomes efficiency as depicted in metric excesses, similar to a bull whip affect in supply chains. This work utilizes a neural network methodology to analyze data describing inpatient cases that incorporates radiology process variables to determine their affect on patient length of stay excesses for a major NJ based healthcare provider. The results indicate that excessive time delays between radiology exam completion times to when results are available to physicians can adversely extend a patient's length of stay beyond initial estimations and imply that enhanced resource allocations at the radiology level may reduce length of stay excesses.

UTILIZING ADVANCED ANALYTICS TO IDENTIFY NETWORK DISRUPTIONS THAT IMPACT PATEINT LENGTH OF STAY METRICS

INTRODUCTION TO HEALTHCARE EFFICIENCY (CLINICAL AND ORGANIZATIONAL PROCESSES)

Robust data resources can provide a vital source to increase operational productivity for healthcare providers. Data repositories containing variables describing various operational attributes for healthcare organizations provide the building blocks to information creation. Variables that describe patient treatment activities, administrative processes, testing procedures and general operational attributes involving various departments of a healthcare organization can provide decision makers with vital information regarding resource effectiveness in patient care.

Productivity gains for healthcare companies can be largely segmented into patient outcome related or organizational process related sectors. Patient treatment and outcomes issues can encompass such factors as diagnosis and treatment effectiveness; while the organizational process sector refers to how efficiently various departmental processes are carried out. Ample research has been conducted on diagnosis and treatment and patient outcome related issues. Studies addressing physician performance [12] and treatment effectiveness [22] are evident, while other studies involving the classification of patient populations into high risk categories for particular ailments are well established [13],[16]. Robust data resources underpinning this sector have enabled researchers to conduct sophisticated empirical testing to test hypothesis.

To more fully address factors that account for efficiencies regarding various performance indicators for health care organizations, analysis incorporating both clinical and process specific related data is required [6]. For example, studies have shown that patients with certain diagnosis may account for high excesses in resource allocations and treatment efficiencies corresponding to established industry metrics, where type of diagnosis or physician performance may be identified as the causal factors driving those excesses. However, the entire process encompassing patient treatment and outcome involves the combination of clinical and organizational procedural related elements [3]. In other words, particular diagnosis and treatment policies of physicians rely on results from lab work, x-rays and departmental processes of the organization, where inefficiencies in any of these processes can result in overall excesses in established metrics such as bed utilization rates or patient length of stay. The entire patient treatment and outcomes process can be compared to that of a supply chain network in non-health care related industries.

The focus of this paper is to incorporate a case study approach to illustrate the effectiveness of data mining analytics to identify factors within the patient treatment and outcomes process that produce high excesses in performance metrics. More specifically it analyzes data describing radiology process subcomponents of the entire patient treatment and outcomes network of activities to determine their impacts on a patient's length of stay for a major NJ healthcare system. The results have indicated that particular radiology process attributes have contributed to inefficient organizational performance as measured by length of stay per inpatient cases. The robust underpinnings of the data mining method (e.g incorporation of algorithmic processing of data) indentified key process inefficiencies and have empowered administrative strategic decision makers to implement new initiatives to address the allocation of resources in this area.

SUPPLY CHAINS AND PATIENT TREATMENT AND OUTCOMES

Traditional supply chain processes involve networks of interrelated components. Product availability at the retail level depends on timely shipment of finished products which depends on efficient manufacturing processes, which in turn relies on efficiency in ordering and shipment of raw materials and parts. A major difficulty in achieving optimal efficiency of an entire supply network revolves around the uncertainty of the performance of each of the components of the network. Some of these uncertainties involve late deliveries, machine breakdowns and order cancellations, to name a few [5]. The end result of a breakdown in any subcomponents of the network is a bullwhip affect or adverse variance in overall outcome performance metrics. By estimating such items as required inventory levels and replenishment times corresponding to expected customer demand at each segment of the network helps reduce inefficient resource allocations down the line of the supply chain [14],[4]. Effective management requires not only concise decision making at the individual task level but must also incorporate the ramifications of decision making across components of the entire process [20].

The entire patient treatment and outcomes process can be compared to that of a traditional supply chain when considering all the variables that comprise the activity. The patient treatment process involves physician treatment activities that depend upon complementary subcomponents of testing procedures and results. Lab tests and radiology exams provide vital information for physicians to administer corresponding treatment policies, where breakdowns or inhibitors of these processes can adversely affect timely treatment activities and corresponding outcome metrics. Given increased data resources in process related activities for healthcare organizations, advanced analytic analysis can be utilized to examine organizational process efficiency for healthcare providers.

INTRODUCTION TO ADVANCED ANALYTICS IN HEALTHCARE

Data Mining Analytics

One way to enhance operational efficiency for healthcare organizations is by more accurately identifying the sources of resource demand and effectiveness of resource allocations of corresponding processes [17],[19]. More formal analytic techniques such as stochastic trees have been utilized to help increase operational efficiencies by enhancing the decision-making process in medical treatment procedures [9,10]. Data mining techniques enable decision makers to identify patterns in clinical, claims and activity-based historical data, to better identify and understand explanatory relationships between variables that describe operational processes [1]. This information can be utilized to devise strategic initiatives to improve organizational performance [23]. Techniques such as clustering, CHAID and CART and neural network methods address segmentation, categorization and predictive capabilities of analytics [7],[8]. The neural network methodology is the utilization of complex computer algorithms that perform multivariate analysis to explain excesses in historical data. One of the advantages of this method is the ability to identify both linear and non-linear patterns that exist among data variables, which is achieved through the incorporation of the Multi-layer Perceptron technique [18]. These advanced analytic methods go beyond mere identification of retrospective capabilities of basic reporting and provide decision markers with quantitative models that describe relationships between variables underpinning processes. These models provide simulation capabilities to project potential outcomes given adjustments to process variable inputs.

The neural network architecture simply creates weights for input data variables to best explain the variance in dependent variables in historical sample data. This weight adjustment process is conducted through the incorporation of a back-propagation testing algorithm [11]. Neural networks are often discussed in context of attempting to mimic the learning and decision-making process of the human brain (e.g., artificial intelligence), since the architecture is composed of interconnected entities, which are compared to neurons of the brain. Neural networks are widely incorporated in the realm of data analysis to enhance predictive capabilities in a variety of business and scientific applications [2],[21].

PROJECT BACKGROUND

Newark Beth Israel Medical Center (NBICM) is an affiliate of the Saint Barnabas Health Care System which is a 670 bed regional referral teaching hospital with specialized programs including heart, lung and kidney transplantation, cardiac surgery oncology and maternal/child health services. The healthcare provider was concerned with excesses in the length of stay for inpatient cases. More specifically, it was interested in better understanding those factors that cause a patient's length of stay to extend beyond that of the pre-designated or estimated length of stay based on industry standards corresponding to patient cases as they are admitted to the hospital. By identifying sources or factors that may lead to increased excesses in length of stay, the organization can more effectively manage operational resources to help mitigate the problem and ultimately reduce costs.

When considering the issue of identifying factors that may lead to a patient's length of stay to extend beyond its diagnosed estimation, initial areas to investigate involve the complexity of a patient's case. Case complexity involves the entire patient treatment and outcomes network of activities, which incorporates such factors as potential for complications, required tests, drugs and procedures, type of physician that is attending, to name a few. However, many of these variables are considered when establishing industry standards according to patient diagnoses and therefore should not significantly account for consistent length of stay excesses. In fact, attending physicians can often be cited as a major factor in high excesses in length of stay, which may actually mask more direct causation of the variance in the performance indicator. More direct causation can be present in organizational resource allocations that can increase throughput time of supportive initiatives to treat a patient. As in supply chain networks or operational Just In Time processes, suboptimal performances of process attributes can increase throughput and corresponding performance metrics down the chain (e.g. time to market for a product or service). For healthcare organizations, this can include elements of the treatment process or issues related to the efficiency of processes in supportive departments that affect a patient's case. The IT department at NBICM decided to focus the analytic initiative on organizational processes, more specifically the efficiency of radiology departmental activities to process examination requests. The network of activities to process an exam was seen as a potential bottle neck that could adversely affect the performance indicator of patient length of stay beyond established standards. The case study involved collaboration with the NBICM's CIO, and IT and data support personnel.

DATA MINING ANALYSIS TO IDENTIFY SOURCES OF LOS (VARIANCE)

In order to address the question at hand data sources needed to be acquired and analyzed. The data utilized for the analysis involved inpatient cases over the years 2006 and 2007. This data source included actual length of stay, industry standard length of stay and general demographic and diagnostic variables according to patient cases. Total data records incorporate some 45,000 in-patient cases. Table (1) provides basic data descriptions.

Insert Table 1 here

Radiology process specific data included variables that described the various activities underpinning the process of conducting various exams (e.g. x-rays) for patients. The entire network of radiology activities corresponding to a patient's exam is described below.

Physicians Request for Radiology Exam
Exam is Scheduled to be Conducted
Exam Begins
Exam is Completed
Exam Results are Authorized to Send to Physician

Table (2) defines the variables that are included in the entire radiology testing process.

Inset Table (2) here

Connecting Radiology Factors to Length of Stay Excesses

In order to link length of stay data to radiology specific activities, inpatient case data sources needed to be appended to radiology specific processes corresponding to patient IDs. The resulting file included patient demographic and diagnostic data (as described in Table 1) with the radiology specific data described in Table 2. The data management process of appending the file revealed some interesting results, namely that patient cases of particular diagnosis' often involved numerous radiology procedures (e.g. numerous radiology exams according to particular diagnosis).

One particular component of interest for NBICM within the network of radiology activities involved the time lag between the completion of an exam to final authorization and receipt of results by a physician, as defined below.

Completed to Finalized Time Gap: Elapsed time from exam completion to authorization of results (measured in hours).

The Completed to Finalized Time Gap variable at the radiology exam level needed to be aggregated to a total Completed to Finalized Time Gap for all radiology exams conducted for particular patient cases. An additional variable (number of exams or radiology procedures according to patient case) was developed as well. These new variables ultimately provided significant explanatory direction of LOS (variance) according to patients.

Data Analytics to identify impacts on LOS

Neural networks are a prominent data mining methodology that incorporates the processing of data with computer algorithms. The mining process identifies consistent patterns and trends in data that explain excesses of performance metrics, which in this analysis, includes variance in LOS, defined below. This approach extends beyond the capabilities of pure sorting and reporting techniques to rank variables that may be associated with levels of performance metrics. It considers frequency and quantitative weighting components to uncover consistency in variable relationships, which ultimately renders the results more robust.

Length of Stay Variance = Actual Patient Length of Stay - Estimated Industry Length of Stay Standard for Corresponding Patient Case: (measured in days).

Modeling Procedures Linking Radiology to LOS (Variance**)**

The modeling scenario incorporated the independent variables of Patient Demographics, Primary Payer, Admit and Discharge Dates, Physician Specialty and detailed radiology procedural variables which now included the Total Hours of Completed to Finalized Time Gap on the dependent variable of length of stay variance per patient case. The Number of Exams variable was then introduced in place of Total Hours of Completed to Finalized Time Gap as an independent variable on the LOS variance in a subsequent analysis. The variables (Total Hours of Completed to Finalized Time Gap and Number of Exams) provided noteworthy impact on a patient's LOS (variance).

Model Results

The neural net model depicted no noteworthy patterns with patient demographics, payer group and physician specialty, however as the total number of hours for all the exams' completion time to finalized times for a particular patient increases, the LOS beyond the estimated or expected LOS per patient increases. In other words, if a patient has multiple procedures or exams that result in aggregate time gaps from completed to finalized activities (e.g. 30 hours), the LOS increases above its estimated time by 2 days and increases as aggregated time gaps increase with a non-linear diminishing positive slope.

This conclusion is further supported by viewing the results of incorporating the Number of Exams variable per patient on LOS (variance). As the number of exams a patient undergoes increases, the LOS beyond their estimated LOS increases. More specifically, this supports the notion that each exam involves a time delay between Completion and Finalized activities that may increase a patient's LOS beyond the estimated LOS, where multiple treatments increase LOS (variance) where the relationship maintains a non-linear positive slope. The data mining results therefore support the notion set out in our hypotheses.

CONTINUED INVESTIGATIVE MINING

In order to more fully leverage the existing data resource, the focus was turned to variables that may impact the Completed to Finalized Time Gap described previously. Data transformations were required in the radiology data file in order to generate more robust variables that could potentially better describe the efficiency of administered exams. Time interval variables in radiology process data in Table 2 (e.g. Completed time) were transformed into two additional variables. These were Day of Week (e.g. Monday, Tuesday, etc) and Time of Day (AM/PM) for corresponding time related variables depicted in Table 3.

Insert Table 3 here

The analytic focus, to identify potential process inhibitors, which result in increased throughput for radiology exams involved incorporating the department where the exam was conducted, type of exam, day and time of exam completion, day and time of exam finalization as independent variables on duration of time in hours between when an exam was completed to when the results were available to the requesting physician as the dependent variable. A neural network methodology was incorporated to analyze the corresponding data. Other investigative models that considered time gaps in other procedural variables (e.g. exam scheduling to exam begin time) were considered as well, but results indicated no noteworthy patterns.

Department Name, Exam Name and most Days of week variables depicted no significant influences on duration of hours between Completed and Finalized exam activities. However, one particular variable consistently registered a significant impact on the Completed to Finalized Time Gap variable. This variable involved the particular time of day when an exam was completed. The pattern identified depicts that when exams were completed in the PM time of day, there was roughly an 11 hour delay to when the exam was finalized and available to the requesting physician. The near ½ day delay could ultimately prove to be a significant driver of a patient's length of stay variance especially when considering multiple radiology exams for particular patient cases.

CONCLUDING REMARKS AND CONTRIBUTIONS TO RESEARCH

The results in this analysis resemble that of a bullwhip affect in a traditional supply chain, where breakdowns in efficiencies within components of the network adversely affect high-level performance metrics. Patient cases that require more radiology exams experience higher length of stay excesses. This was supported by the advanced analytics, which identified that process inhibitors in radiology exam activities accounted for increases in patient lengths of stay. The process inhibitor in this case was idle time that existed from when an exam was completed to when it was authorized and available for the attending physician to review. More specifically, when exams were completed after noon, there was a consistent lag to when results reached attending physicians. Given this information, decision makers can more concisely review radiology resource allocations to address the problem. This could involve review of staffing allocations corresponding to radiology exam supply and demand factors. Reducing the idle time between exam completion and finalization subcomponents could reduce the time that attending physicians wait to apply corresponding treatments given review of test results and therefore mitigate LOS variance.

The incorporation of robust quantitative analytic techniques also provide analysts with greater legitimacy in proposing strategic initiatives which supports earlier research addressing the principal agency relationship in health care systems [12]. Analysis' based in quantitative and statistical methods that identify consistent patterns and trends in data resources provide legitimacy in reliability of results and therefore can enhance the implementation process of strategic initiatives. This concept is particularly noteworthy given the sometimes conflicting focus between physician treatment activities and organizational efficiency goals. In order to enhance the efficiency of organizational process outcomes, administrators can communicate results of robust analytics to physicians involved in patient treatment activities to achieve greater cohesion in considering the most appropriate methods that achieve the best care for patients along with organizational productivity.

Length of stay proves to be a critical metric for health care organizations and the complexity of fully identifying all factors that may impact this variable remains. Applying concepts established in supply chain optimization research can provide the fundamental base to addressing this issue. In this case, the subcomponents of radiology activities supporting physician treatment activities, illustrates this concept. Future research could involve a more detailed analysis of the subcomponents that comprise networks of activities including other functional areas of a healthcare organization (e.g. ER, Lab Testing). The radiology network includes subcomponents involving time gaps between ordering and scheduling exams, scheduling to beginning exams, beginning to complete exams and completing to finalizing. Inefficiencies in any of these could adversely affect higher level performance metrics for the healthcare provider. In order to more fully understand excesses in performance metrics, organizations must address the various networks of activities (administrative, departmental, physician related) that underpin the hospital patient care process to examine the interdependencies of elements that ultimately impact performance metrics. The combination of data resources and advanced analytics can ultimately provide decision makers with robust models that identify bottlenecks in the network of activities involved in treating a patient. Corresponding strategic initiatives enhancing resource allocations to address process throughput and bottlenecks can increase organizational productivity.

References

[1] Bar-Or, A., Keren, D., Schuster, A. and Wolff, R. (2005), "Hierarchical Decision Tree Induction in Distributed Genomic Databases, IEEE Transactions on Knowledge & Data Engineering, (17:8), pp. 1138-1151.

[2] Bishop, C. 2005, Neural Networks for pattern recognition. Oxford University Press.

[3] Cheglowski, R., Churilov, L. and Wasserthiel, J. (2007), Combining Data Mining and Discrete Event Simulation for a value-added view of a hospital emergency department, Journal of Operational Research Society, (58:2), pp. 246-254.

[4] Chen, F., Drezner, Z., Ryan, J. K., & Simchi-Levi, D. (2000), Quantifying the bullwhip effect in a supply chain: The impact of forecasting, lead times, and information, Management Science, 46(11).

[5] Davis, T. (1993), Effective Supply Chain Management, Sloan Management Review, 34(4).

[6] Goodall, Colin. (2009), Data Mining of Massive Datasets in Healthcare, Journal of Computational & Graphical Statistics, (8:3), pp. 620- 635.

[7] Han, J., and Kamber, M.(2001), Data Mining: Concepts and Techniques, Morgan Kaufmann, San Francisco.

[8] Hand, D., Mannila, H., and Smyth, P. (2001), Principles of Data Mining (Adaptive Computation and Machine Learning**),** MIT Press.

[9] Hazen, Gordon. (2000), "Preference Factoring for Stochastic Trees," Management Science, 46(3), pp 389-403.

[10] Hazen, G. (1992), Stochastic Trees: A new technique for temporal medical decision modeling. Medical Decision Making, (13), pp. 133-119.

[11] Hinton, G. (1992), "How Neural Networks Learn from Experience", Scientific American.

[12] Kohli, R., and Kettinger, W. (2004), Informating the clan: Controlling Physicians' Costs and Outcomes", MIS Quarterly, 28(3), pp 363-394.

[13] Kudyba, S., Hamar, B., and Gandy, W. (2005), Enhancing efficiency in the healthcare Industry: Using critical data and predictive modeling can help identify high-risk candidates in a total health plan population, Communications of the ACM, (28:12), pp. 107-110.

[14] Lee, H., Padmanabhan, V. and Whang, S. (1997), The bullwhip effect in supply chains, Sloan Management Review, 38(3).

[15] Lee, H., Kut, C. and Tang, C. (2000), The Value of Information Sharing in a Two-Level Supply Chain. Management Science., 46(5).

[16] Lieu, T., C. Quesenberry, M. Sorel, G. Mendoza, and A. Leong. (1998), "Computer-based models to identify high-risk children with asthma," American Journal of Respiratory and Critical Care Medicine, 157(4), pp.1173-1180.

[17] McLoughlin C, S., Yang and R. Van Dierdonck. (1995), Professional Service Organizations and Focus, Management Science, 14(7), pp. 1185-1193.

[18] Rumelhart, D. and McClelland, J. (1996), Learning Internal Representations by Error Propagation, MIT Press.

[19] Smith-Daniels, V., Schweikhart, S., & Smith-Daniels, D. (1988), Capacity Management in health care services: Review and future research directions, Decision Sciences (19), pp. 889-919.

[20] Stevens, G.C. (1990), Successful Supply-Chain Management. Management Decision, (28:8).

[21] Swingler, K. (1996), Applying neural networks: A practical guide. Academic Press, New York.

[22] Veluswamy, R. (2008), Golden Nuggets: Clinical Quality Data Mining in Acute Care, Physician Executive, (34:3), pp. 48-53.

[23] Xiaohua, H.., Il-Yeol, Song., Hyoil, Han. (2005), Temporal Rule Induction for clinical outcome analysis, International Journal of Business Intelligence and Data Mining, 1(1), pp. 122-136.

Tables 1 - 3

Table (1) Length of Stay Data File

Patient ID	Patient Identifier
Gender	Patient Gender
Age	Patient Age
Primary Payer	Insurance Provider Code
Diagnostic Code	Principal Diagnosis Identifier for Patient
Length of Stay	Actual Length of Stay at Hospital
Length of Stay	Length of stay according to diagnosis/industry standard
Admittance Date	Date of Patient admittance
Discharge Date	Date of Discharge
Physician Specialty	Attending Physician Specialty

N = 45,000

Table (2)

Radiology (Descriptive Procedural Characteristics)

Name of Department	(Department where exam is conducted)
Name of Exam	(Type of exam requested)
Schedule Time	(Scheduled time to begin an exam)
Begin Time	(Actual time exam was conducted)
Completed Time	(Time of completion of exam)
Finalized Time	(Time when exam results were authorized and sent to physician)

N = 120,000

Table 3

Detailed Time-Based Procedural Variables included:

Completed Day	(Day of week exam was completed)
Completed AM/PM	(Time of day exam was completed)
Finalized Day	(Day of week results of exam were authorized)
Finalized AM/PM	(Time of day results of exam were authorized)

DETECTING CARGOES WITH POTENTIAL SECURITY FRAUD – THE APPLICATION OF SELF-ORGANIZING MAP NETWORKS

Melody Kiang, Information Systems Department, California State University, 1250 Bellflower Blvd, Long Beach, CA 90840, (562) 985-8944, mkiang@csulb.edu
Robert Chi, Information Systems Department, California State University, 1250 Bellflower Blvd, Long Beach, CA 90840, (562) 985-4238, rchi@csulb.edu

ABSTRACT

Kohonen's Self-Organizing Map (SOM) network is one of the most important network architectures developed during the 1980's. The SOM network was originally designed for solving problems that involve tasks such as clustering, visualization, and abstraction. In this research, we evaluate the feasibility of applying SOM networks to build a cargo security warning system that can identify cargoes with higher potential of security fraud.
Keywords: SOM network, neural networks, cargo security, early warning system.

INTRODUCTION

For many centuries researchers have been trying to build intelligent computer systems that can perform daily routine of mankind for the purpose of replacing human from tedious tasks or hazardous work environment. To make intelligent computer systems requires an understanding of how human process information. The inspiration for neural network architecture originally came from studies of biological nervous systems. People have long recognized that human beings and animals are much better and faster at processing and recognizing speech and images than most sophisticated computers. The field of artificial neural networks (ANNs) began as an approach to imitating human intelligence for building artificial intelligence systems that can learn from experience.

ANNs with their remarkable computational power can extract information and uncover meaningful trends from complicated and massive amount of data that are too complex to be processed by humans or other conventional computer techniques. For example, credit card companies are also able to prevent fraudulent activity on their customer's credit cards by using ANNs to map and predict a customer's purchase behavior and flag activities that are determined to be "suspicious" in nature by the ANNs. Today, many financial service providers in the US including American Express, Washington Mutual Bank, Wells Fargo and MasterCard use ANN's. Even though success metrics are hard to come by, industry analysts all agree that the use of ANN's has enabled these firms to efficiently and effectively cut down on fraud activity in turn realizing significant savings from losses of covering fraud charges on customer accounts [1].

Every year, more than 5,500 vessels carrying 4.5 million cargo containers pass through airports and harbors of the U.S. With $200 billion worth of goods flowing in and out of Los Angeles and Long Beach, these ports have a large influence on the economy, across several different industries including aerospace, electrical, consumer products, textiles, among others [2]. Given the importance of the ports and their significant influence on the urban population, it becomes more and more important that some type of real-time monitoring/early warning system is needed to ensure the security and safety of the ports. In this research, we apply Kohonen's SOM network to identify cargoes with higher potential of security fraud.

ARTIFICIAL INTELLIGENCE IN PORT SECURITY

Ever since the terrorist attacks of September 11, 2001, security has become a national epidemic. Great emphasis has been placed on securing the U.S. borders to help minimize any threat of attack. With that, it has been increasingly recognized that the country's ports represent a significant point of entry and vulnerability. For example, in 2004, customs inspectors started scanning more than double the percentage of packages compared with 2003 after recognizing the potential threat. Nevertheless, only 5.4 percent of all incoming cargo is scanned for contraband given the enormous cost and time necessary for manual inspection. Even with the advent of gamma, X-ray, and radiation detection technology, manual inspection of all cargo coming from just the ports of Los Angeles and Long Beach alone would not be practical given the processing time and financial implications, and the fact that these two ports represent 30 percent of all U.S. international sea trade.

Known that the current security controls in place at the ports are not adequate for today's environment, an information system solution to support safe movement of cargo would greatly improve safety with a minimum impact on current processes. According to Rod MacDonald, acting assistant commissioner of the U.S. Customs Border Protection's Office of Information and Technology, collecting and reviewing shipment information before arrival is the best solution to combating illegitimate cargo. Application of an information technology solution that takes into account the variability of the cargo process and container variations would be ideal. Therefore, this research explores the possibility of applying artificial intelligence (AI) and machine learning techniques such as neural networks to the security issue at hand. Information is one of the most valuable assets for any organization, and neural networks can exploit this advantage through in-depth analysis and informed decision-making.

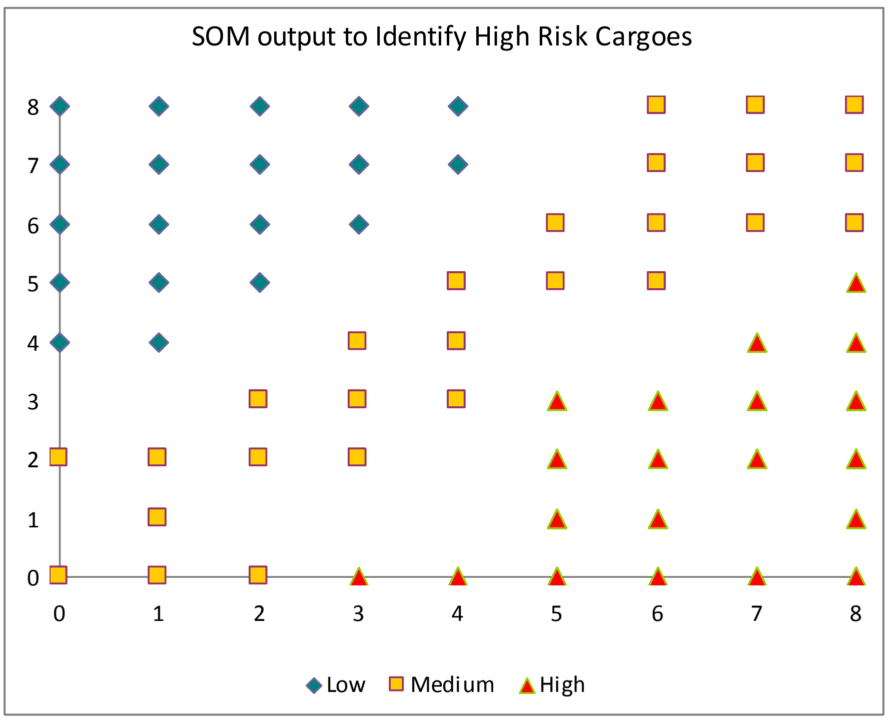
SOM Network is a special type of neural network that can learn from complex, multi-dimensional data and transform them into visually decipherable clusters. The theory of the SOM network is motivated by the observation of the operation of the brain. Various human sensory impressions are neurologically mapped into the brain such that spatial or other relations among stimuli correspond to spatial relations among the neurons organized into a two-dimensional map [3]. The main function of SOM networks is to map the input data from an n-dimensional space to a lower dimensional (usually one or two-dimensional) plot while maintaining the original topological relations. The physical locations of points on the map show the relative similarity between the points in the multi-dimensional space. In other words, the data points that were close in the higher dimensional space should remain close in the reduced lower dimensional map. Therefore, when grouped into clusters, the cargoes that possess similar attributes values will be in the same cluster. Besides providing the cluster membership information, the SOM visual map clearly depicts the actual relationship among the cargoes within and among different risk groups.

The objective of this study is to explore the possibility of applying SOM neural networks as an early warning system to alert port authorities with cargoes of potential security fraud. When fed the appropriate data inputs regarding cargo containers, this system would "learn" the differences between potentially lawful and unlawful cargoes. Unlawful cargo would include forms of contraband such as untaxed cigarettes, infested fruit, counterfeit software, illegal immigrants, narcotics, drugs, "dirty bombs" (i.e., explosives filled with nuclear waste), weapons, and other terrorist devices. X-ray, radiation scanners, and cameras can only go so far as to detect these forms of contraband. A neural networks based early warning system will allow for port personnel to focus their investigation efforts on cargoes that are more likely to contain contraband, which helps to improve detection efficiency and safety.

EXPERIMENTAL DESIGN AND PRELIMINARY RESULTS

Because this study is a preliminary investigation of the applicability of SOM to cargo security-warning system, we first focused on artificially constructed data. Knowledge of the true data generating mechanisms (the correct cluster membership for each observation) is essential for valid comparisons with respect to the accuracy with which SOM recovers the true cluster structures. A total of 150 random observations were generated that contains 5 observed variables, representing the five important factors: 1) port and country of origin, 2) shipping line/forwarder, type, 3) shipper/exporter and consignee, 4) nature of goods on manifest, and 5) nature of paperwork files/filled out, were equally distributed among the three segments (clusters), low, medium, and high risk. The SOM output map is shown is Figure 1.

FIGURE 1
SOM Output of the Simulated Data Set



Predictive models based on statistical techniques have been widely adopted and some techniques perform reasonably well in terms of rate of correctness in their predictions. However, all methods require complex and advanced analytical skill to explain and interpret the output from the prediction model. In this study, we introduce the Self-Organizing Map (SOM) Network that can learn from complex, multi-dimensional data and transform them into visually decipherable clusters on an output map. A salient feature of using SOM method over the other approaches is the added visual map. The 2-dimensional plot provides an easy-to-read graphical interface that does not require specialized analytical knowledge to interpret the results. It is a valuable decision support tool that helps the port personnel visualizes the relationships among inputs. Therefore, the port personnel can interactively determine the composition of the clusters using the output map of SOM and incorporate subjective criteria when desired.

ACKNOWLEDGEMENT

This research was entirely supported by Metrans Center and its funding agencies: the Department of Transportation and California Department of Transportation. The opinions are of the authors'.

REFERENCES

[1] Balint, Kathryn, "Fraud fighters, Fair Isaac Corp. products protect consumers, companies," SignOnSanDiego.com, the Union-Tribune, Feb. 18, 2005.

[2] Grossberg, J.," No safe harbor for smugglers," Daily Breeze: A1, A8, April 9, 2006.

[3] Kohonen, T., *Self-Organizing Maps*, Springer, 1995.

TEAM COGNITION IN TECHNOLOGY-MEDIATED VIRTUAL TEAMS

Hayward Andres, School of Business & Economics, North Carolina A&T State University,
1601 East Market Street, Greensboro, NC, 27411, hpandres@ncat.edu

ABSTRACT

Social impact theory and group action theory are integrated to provide a framework for explaining how technology-mediated collaboration constrains or enhances team shared mental model development. Social impact theory and group action theory are then used to explain the role of shared mental models in shaping team work quality that ultimately affects both task productivity and team process satisfaction outcomes. Partial least squares analysis revealed that technology-mediated collaboration does impact shared mental model development. Further, the results also demonstrate that shared mental model helps to facilitate teamwork quality needed for successful task productivity and team process satisfaction outcomes.

INTRODUCTION

Companies are increasingly turning to a new business model, the virtual organization, where virtual teams are constructed with individuals that 1) are geographical dispersed, 2) are linked via collaboration technologies, and 3) collaborate across time and space in synchronous and/or asynchronous mode. To better understand technology-mediated collaborative problem solving, a direct observational research approach is used in examining the mediating effects of shared mental model and teamwork quality on task outcomes.

LITERATURE REVIEW

Team-based collaborative problem solving is a social process dependent on effective and efficient information exchange and interpretation. Cognition in both individuals and teams involves the acquisition, exchange and sense-making of information to construct a mental model. The extent of accuracy or fit of a team member's mental model with a specific task situation can influence task participation rates and performance. Specifically, shared mental models have been shown to facilitate improved performance by virtue of accurate anticipation of appropriate contribution, team member needs or expectations, and reduced communication overhead [9]. Social Impact Theory (SIT) suggests that changes in feelings, motivations, and behavior occur in an individual as a result of the real, implied, or imagined presence or actions of other individuals involved in a social interaction [10]. Social impact theory (SIT) suggests that behavior is guided by social influence derived from 1) salience or importance attributed to team members (*strength*), 2) time, spatial, or interpersonal distance among team members (*immediacy*), and 3) the quantity of influential sources (*numbers*). Hacker's [6] action regulation theory (ART) provides a framework for explaining how work behavior is internally guided and regulated. According to the ART framework, there are three mechanisms that make up action regulation - 1) orientation to the task, 2) solution implementation and 3) evaluation. *Orientation* focuses on assessing task requirements, goal setting, and defining strategies for goal attainment. The *implementation* process is guided by continuous feedback on team member contribution toward goal achievement, and is followed by *evaluation* of the current form of the task solution. In other words, the quality of work-related actions are controlled by goals derived from interpretation (i.e. mental model) of task requirements and status and these goals undergo *redefinition* based on evaluations of work completed.

RESEARCH MODEL AND HYPOTHESES

The research model is depicted in Figure 1 below. The model suggests that collaboration mode will impact shared mental model development. Moreover, team work quality is a function of the accuracy of and scope of a team's shared mental model of the task requirements and appropriate task execution procedures. Finally, teamwork quality impacts task productivity and team process satisfaction outcomes.

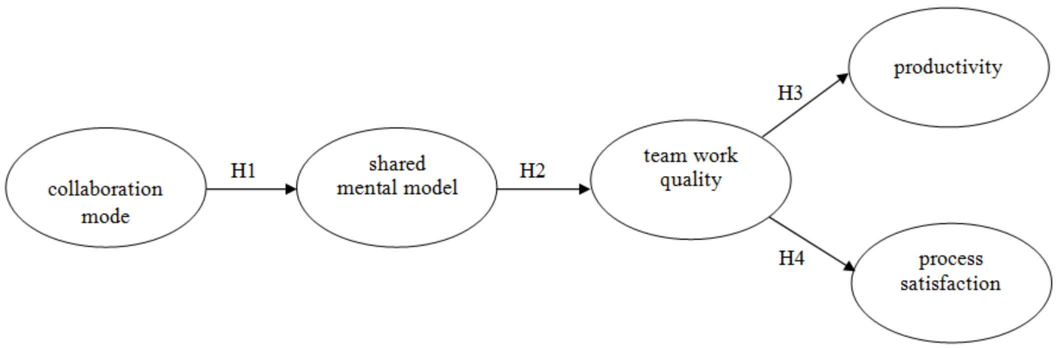


Figure 1. Research Model

Collaboration Mode and Shared Mental Model

Dennis et al. [5] noted that technology-mediated collaboration can be inferior to face-to-face collaboration because of its ability to facilitate the construction of verbal information or messages that are supplemented with physical gestures or nonverbal cues (e.g., postures, facial expression, eye gaze, tone of voice, and conversation pauses). These nonverbal cues function as feedback that confirms or disconfirms understanding and controls turn-taking. Further, the coordinated patterns of communication facilitate immediate feedback which minimizes potential disengagement thereby maintaining a shared focus that allows the team to move forward with minimal delays and minimal instances of deadlock or impasse. SIT suggests that lack of team member proximity associated with technology-mediated collaboration can limit motivation for team member participation needed to offer and evaluate ideas that could be used by the group in shared mental model development. The diminished motivation can be attributed to failure to attribute salience or credibility to contributions from distant team members (via lower experienced *strength*, *immediacy* and *number* effects) and through efficacy-based disengagement tendencies [2] [6]. Thus the following hypothesis is proposed.

HYPOTHESIS 1. *Groups collaborating in the face-to-face settings will develop a more accurate shared mental model of the task requirements and task status than the technology-mediated groups.*

Shared Mental Model and Teamwork Quality

Bonito [3] showed that when the degree of shared mental model was low among team members, contribution and participation was dominated by those with more shared knowledge. Alternatively, a high degree of shared mental model similarity resulted in greater team-wide contribution and discussions of alternative solutions. Group action regulation theory suggests that the key mechanism by which a shared mental model helps to facilitate coordinated action is via shared or common *redefinitions* of the problem and consistent *evaluations* of the appropriateness of the current solution and strategy [6]. Finally, social impact theory suggests that as the number of team members with shared similar opinions and ideas increase (via strength and number effects), others are influenced to seek and attain convergence by submitting contributions complementary and synergistic to the current majority-held shared mental model [2]. Thus, the following hypothesis is proposed.

HYPOTHESIS 2. *Greater accurate shared mental model development will be associated with greater team work quality.*

Teamwork Quality and Task Outcomes

A shared mental model helps team members coordinate individual contributions and use anticipation and prediction to minimize duplication of effort. Using implicit coordination which is enacted tacitly through anticipation and adjustment based on shared mental models, team members can work in a coordinated manner without the need for extensive communication. In addition, maintenance of an encouraging and supportive climate has been shown to motivate participation and create a sense of team efficacy that then leads to improved task outcomes [7]. Cooperative team level goals and a positive team climate have also been shown to result in increased team member satisfaction with both interpersonal interactions and task solution [7, 11]. Hence, the following hypotheses are proposed.

HYPOTHESIS 3. *Greater teamwork quality will be associated with greater team productivity.*

HYPOTHESIS 4. *Greater teamwork quality will be associated with greater team process satisfaction.*

RESEARCH METHODOLOGY

To test the research model and hypotheses, a laboratory experiment was conducted to examine the effects of two different modes of team collaboration – face-to-face and non-collocated technology-mediated collaboration. The participants were forty-eight Management Information Systems undergraduate students familiar with the Systems Development Life Cycle approach to software design and knowledge of structured programming. The teams, comprised of four members, were required to enhance the functionality of a hypothetical university information system. The experimental task required each team to construct software design documentation that included (1) a hierarchy chart, (2) a list of function prototypes, and (3) pseudocode for each function identified as part of a solution to the problem.

Measures

In providing their observation ratings of shared mental model and teamwork quality, three trained observers used a rating scale that ranged from 1 (very low) to 7 (very high). The interrater agreement index for all scale ratings ranged from $r_{wg(j)} = 0.72$ to $r_{wg(j)} = 0.99$ indicating very good interrater agreement [8]. Scale items appear in Table 1 below. The team productivity measure was determined by awarding one point for each correct specification of any data value needed in a specific data file, correct output and input data value of a program module (i.e., function or subroutine), and correct specification of a program statement needed in a program module. Team process satisfaction was assessed by administering a questionnaire after task completion.

Measurement Model

Composite reliability scores for every construct (ranging from 0.873 to 0.970, see Table 1) are well above 0.70, which is the suggested benchmark for acceptable reliability [4]. In addition, the t-statistics for the item to construct loadings were all significant at $p \leq .01$. These results indicate that the measurement model has displayed both item internal consistency reliability and item convergent validity.

Table 1. Composite Reliability, AVE, and Indicator Loadings

Construct and Item Level Values		**loading**
Shared Mental Model (Composite Reliability = 0.970; AVE = 0.889)		
smm1	It was clear that everyone developed the same level of understanding of the task requirements.	0.930
smm2	There was some difference of opinion or concern about the correctness of the proposed solution.	0.963
smm3	There was significant confusion about what was going on.	0.941
smm4	Some team members required a lot of explanations about what was going on.	0.936
Teamwork Quality (Composite Reliability = 0.919; AVE = 0.741)		
tmwk1	Team-wide consensus was confirmed before moving forward with an idea.	0.925
tmwk2	Ideas were thoroughly discussed and evaluated among all team members.	0.901
tmwk3	Team made an obvious effort to create and maintain a positive climate.	0.747
tmwk4	Team would clearly enjoy working together on another project.	0.859
Satisfaction (Composite Reliability = 0.873; AVE = 0.633)		
satisfac1	To what extent did you feel frustrated or tense about another team members' behavior?	0.695
satisfac2	How satisfied were you with the quality of your team's solution?	0.796
satisfac3	To what extent do you feel committed to the team's solution?	0.838
satisfac4	To what extent do you feel confident that the team solution is correct?	0.845

Discriminant validity is evidenced when all the loadings of the scale items on their assigned latent variables or construct are larger than their loading on any other latent variable. Table 2 below provides the correlations of each item to its intended latent variable (i.e., loadings) and to all other constructs (i.e., cross loadings).

Table 2. Indicator Loadings

	Latent Variable Item Loadings		
Indicator	**shared mental model**	**team work quality**	**satisfaction**
smm1	**.930**	.856	.432
smm2	**.963**	.798	.323
smm3	**.941**	.579	.037
smm4	**.937**	.597	.060
tmwk1	.729	**.925**	.662
tmwk2	.757	**.901**	.540
tmwk3	.532	**.747**	.544
tmwk4	.560	**.859**	.621
satisfac1	.521	.617	**.695**
satisfac2	-.075	.396	**.796**
satisfac3	.229	.646	**.845**
satisfac4	-.125	.388	**.838**

Table 3 below indicates that the AVE square roots that appear in the diagonal are larger than any correlation between the associated construct and any other construct [4]. This AVE analysis result and the item to construct loadings discussed above suggest that the measurement model displays discriminant validity.

Table 3. Latent Variable correlations and square root of AVE

	shared mental model	**team work quality**	**satisfaction**
shared mental model	**0.943**		
team work quality	0.762	**0.861**	
satisfaction	0.242	0.684	**0.796**

Note: square root of the constructs' AVE appear in the diagonal

Structural Model

Using PLS Graph (Version 3.0 Build 1130), the structural model and hypotheses were assessed by examining path coefficients and their significance levels [4].

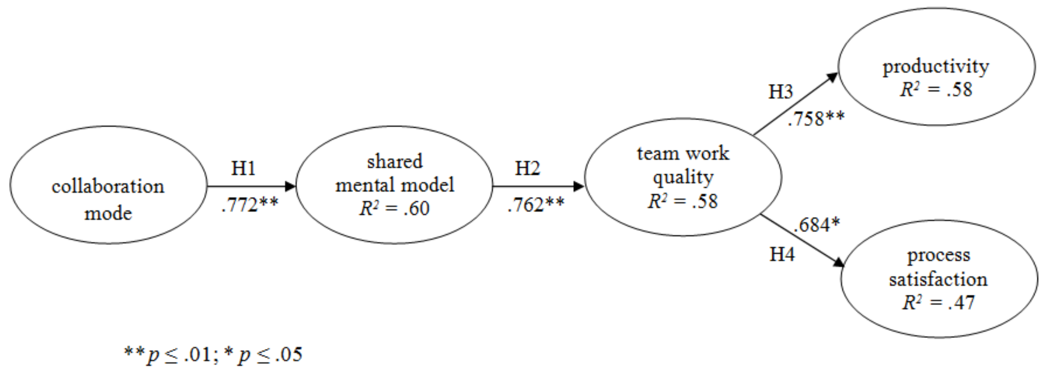


Figure 2. PLS Analysis Results

The PLS analysis results (Figure 2) show that all the hypotheses were supported. Collaboration mode was shown to increase shared mental model (b = 0.772, t = 6.230, $p \leq .01$) thereby supporting hypothesis 1. Hypothesis 2 was supported in that shared mental model increased teamwork quality (b = 0.762, t = 12.336, $p \leq .01$). Teamwork quality lead to increases in both productivity (b = 0.758, t = 10.875, $p \leq .01$) and team process satisfaction (b = 0.684, t = 2.625, $p \leq .05$) indicating support for hypothesis 3 and hypothesis 4 respectively. All of the mediation paths in the model were shown to conform to relevant requirements indicating that 1) team reflexivity mediated team learning on productivity and interaction quality and 2) team learning mediated the impact of collaboration mode on team reflexivity [1].

DISCUSSION/CONCLUSION

By integrating social impact and action regulation theories, insights and theoretical contributions that emerge from this study relate to how and why (1) technology-mediated collaboration can diminish participation; (2) collaboration mode can dictate which ideas get incorporated into a team's shared mental model; and (3) shared mental model can regulate both task-related and social-related behaviors.

REFERENCES

[1] Baron, R. M., and Kenny, D. A. "The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations." *Journal of Personality and Social Psychology*, 1986, 51(6), 1173-1182.

[2] Blaskovich, J. L. "Exploring the effect of distance: An experimental investigation of virtual collaboration, social loafing, and group decisions." *Journal of Information Systems*, 2008, 22(1), 27-46.

[3] Bonito, J. A. "Shared cognition and participation in small groups: Similarity of member prototypes." *Communication Research*, 2004, 31(6), 704-730.

[4] Chin, W. W. "Issues and opinions on structural equation modeling." *MIS Quarterly*, 1998, 22(1), 7-16.

[5] Dennis, A., Fuller, R. M., and Valacich, J. S. “Media, tasks and communication processes: A theory of media synchronicity.” *MIS Quarterly*, 2008, 32(3), 575-600.

[6] Hacker, W. “Action regulation theory: A practical tool for the design of modern work processes?” *European Journal of Work and Organizational Psychology*, 2003, 12(2), 105-130.

[7] Hoegl, M., Weinkauf, K. H., and Gemuenden, G. “Interteam coordination, project commitment, and teamwork in multiteam R&D projects: A longitudinal study.” *Organization Science,* 2004, 5(1), 38-55.

[8] James, L. R., Demaree, R. G., and Wolf, G. “Estimating within-group interrater reliability with and without response bias.” *Journal of Applied Psychology*, 1984, 69(1), 85-98.

[9] Klimoski, R., and Mohammed, S. “Team mental model: Construct or metaphor?” *Journal of Management*, 1994, 20(2), 403-437.

[10] Latane, B. “The psychology of social impact.” *American Psychology,* 1981, 36(4), 343-356.

[11] Tjosvold, D., Yu, Z., and Chun, H. “Team learning from mistakes: The contribution of cooperative goals and problem-solving.” *Journal of Management Studies,* 2004, 41(7), 1223-1245.

SOME PHILOSOPHICAL THOUGHTS ON MACHINE CONSCIOUSNESS

Jinchang Wang
School of Business, Richard Stockton College of New Jersey, Pomona NJ 08215
jinchang.wang@stockton.edu

ABSTRACT

We present our philosophical thoughts on the perspectives of development of computer intelligence and computer consciousness, such as difference between intelligence and consciousness, how intelligent / conscious a machine can be, the inspiration of the Godel's incompleteness theorem, Searle's Chinese Room argument and the Turing Test, and logical difficulty in "mind-copying".

1. CONSCIOUSNESS VS. INTELLIGENCE

It is not easy to separate intelligence from consciousness completely. Some consciousnesses are parts of intelligence or closely related to intelligence, such as inspiration, intuition, creativity, and originality. But intelligence in most cases refers to those brain activities of making decisions, learning knowledge and skills, understanding / creating knowledge, and developing / using technology and tools. Consciousness, on the other hand, refers to those brain activities of mental reactions to the outside world. Intelligence reflects the capacity of thinking. Consciousness reflects the capacity of feeling.

Question "whether a computer can be more intelligent than humans" is viewed as a well-posed question, while for question "whether a computer can be more conscious than humans", we would take it as an ill-posed question because we do not know what "more conscious that humans" means exactly. Question "whether a computer can be as conscious as a human" makes sense.

2. HOW INTELLIGENT / CONSCIOUS A MACHINE CAN BE

People have widely different opinions on these two questions: - Can a computer eventually be more intelligent than humans? Can a computer be as conscious as humans?

Strong AI believers say: Yes. They claim that human's intelligence can be surpassed by a machine and a full range of consciousness can be realized on a machine if computer programs are sophisticated sufficiently [1][4][5][6][7]. "Within several decades information-based technologies will encompass all human knowledge and proficiency, ultimately including the pattern-recognition powers, problem-solving skills, and emotional and moral intelligence of the human brain itself." [3] They argued that computer's capability of thinking is enhanced continuously and exponentially, albeit its current level is still very low, while human's intelligence has remained almost unchanged for thousands of years. Kursweil, an optimistic futurist, predicted, "By the late 2020s, we will have completed the reverse engineering of the human brain, which will enable us to create nonbiological systems that match and exceed the complexity and subtlety of humans, including our emotional intelligence."[3] " I set the date for the Singulariy – representing a profound and disruptive transformation in human capability – as 2045. The nonbiological intelligence created in that year will be one billion times more powerful than all human intelligence today." [3] But these are his just "beliefs" without being supported by compelling arguments.

Dualists and theists, on the other hand, say: No. They assert that intelligence, consciousness and spirit are the unique gifts from the God and are something special which are not material, and it is not possible to realize them by using the materials in this world. But facing the reality that current computers are already able to do something intelligent, such as solving mathematical problems and language translation, they failed to indicate what is the critical line that computers can never reach or surpass.

Many scientists and philosophers are wary on machine consciousness and machine intelligence. Searle used a case called Chinese Room to show that a machine could never be as intelligent as humans because it did not understand what it was doing. [13] Hofsdadter contemplatively argued that it was self-reference that made human's intelligence and consciousness complicated and subtle, which might not be completely solvable [4]. Penrose [8][9][10] showed that some human's intelligence was not algorithmic therefore was not able to be coded into a computer. He argued that quantum effects of indetermination were present in the human brain and could be responsible for non-algorithmic functions, which did not show up in a digital system.

It is an undeniable fact that computer's intelligence is improved with a fast, and increasingly fast, rate, while human's intelligence stays. It is a reasonable prediction that computer will catch up with human in terms of intelligence sooner or later, unless we can identify some points which a computer is not able to reach at all. That is, a machine can be eventually more intelligent than, and as conscious as, humans, unless there exist some points that are proved to be unreachable by a machine. Here we talk about whether it is possible, rather than whether it is probable. Showing the technical complexity and difficulties would be an argument of "improbableness", rather than a proof of "impossibleness". The proof of "impossibleness" would come from logic, showing a contradiction would inevitably occur if some point of intelligence is reached by a machine. As far as one has not yet identified such a "forbidden point", the viewpoint of strong AI on machine intelligence is not refutable, and we could say only that it seems so far nothing could stop a computer to be more intelligent than humans eventually.

As some scientists and philosophers alluded to, self-consciousness or self-identity was a point that is not only difficult but likely impossible for a machine to realize. But so far no hard proof has been given.

3. INSPIRATIONS FROM GODEL'S INCOMPLETENESS THEOREM

Godel's incompleteness theorem indicates the limitation of doing something within a closed system, which implies a warning to those who are thinking of man-made intelligence and man-made consciousness [2].

Godel's theorem shows that there are true statements in a mathematical system which are not provable as the theorems of the system. It implies that a closed system is not capable of exhausting all the truths of that system without the help from outside of the system. That is not due to the huge number of such statements. It is because that some true statements cannot be derived at all if one stays inside of the system.

Godel's incompleteness theorem advises all scientific researchers to be humble. Staying inside this world, there must be something in this world we can never comprehend.

We believe the power of science and logic in the process of trying to understand the world around us. As AI researchers in automated reasoning for twenty years, we applaud the achievements in making computer more intelligent, and we realize that many AI accomplishments have been achieved thanks to strong and optimistic confidence on the 'omnipotent' computers. There is a huge potential for the development of AI. On the other hand, as scientific and philosophical thinkers, we doubt that science is able to resolve all problems related to ourselves, such as intelligence and consciousness. Research on human's intelligence and consciousness is researcher's investigating himself, which is very different from the other researches that are researcher's investigating something else.

4. SEARLE'S CHINESE ROOM AND THE TURING TEST

John Searle says, "Someone is bound to ask, can you prove that the computer is not conscious? The answer to this question is: Of course not. I cannot prove that the computer is not conscious, any more than I can prove that the chair I am sitting on is not conscious." [13] A computer can have some consciousnesses of humans, or may have most consciousnesses of humans, but computers cannot have the full range of consciousnesses of humans, since there is some consciousness that can never be possessed by computers. Searle argued by using his Chinese Room that a machine may look to have consciousness but it is not really conscious [13], which started a hot and up-to-date debate. Put aside the issues of what the real consciousness is and how to check it, statement "a computer cannot be as conscious as human" could be a true statement any way.

It seems there is no solid reason to deny that a machine can eventually pass the Turing Test [12], in terms of either intelligence or consciousness. A machine that passed Turing Intelligence Test can be as smart, and as stupid, as humans. If it were proved that some consciousness of humans cannot be realized on a machine, then a machine that passed Turing Consciousness Test would not be as conscious as human. It must have passed Turing Consciousness Test by cheating. It would look as conscious as humans, but it would not have the full range of human's consciousness. Note that we cannot say that a machine having passed the Turing Intelligence Test is not as intelligent as humans, even though the "way of thinking" of the machine is different from that of ours. Intelligence can be measured and evaluated from the outside of the intelligence carrier; while consciousness contains something deep in the consciousness carrier which cannot be measured from the outside. In this sense, we say Turing Consciousness Test is incomplete, while Turing Intelligence Test is complete. That is the difference between Turing Intelligence Test and Turing consciousness test.

5. LOGICAL DIFFICULTY IN "MIND-COPYING"

Kurzweil depicts a scene of "mind-copying" or reverse engineering [3][4][11]: - A person's brain is scanned so that it can be stored, copied, rebuilt, and transported in light speed. If it were possible, then a person would be immortal.

But there is a logical difficulty on copying "self". Kurzweil asserted that if someone's entire memory, way of thinking, habits, consciousnesses, and whatever one can think about the "self", are scanned and copied, then 'that oneself' is copied. Is the copy is exactly the same self as original one? Suppose we have such a copy of self. The original self is identical to the copy of it in terms of original self's past. But how about the future of them? Even though we may have two identical "past self's", we cannot have two identical "future self's". That is because each of the two "past self's" has its own free will and moves freely, they will have different experiences from the time the copy is made, unless they move together and think identical (since they were built identical inherently). But even they move together and think identically, how do they deal with each other? If one is hit, how does the other react? If they do not move and think together, each self would have a location at any moment and have a particular situation to deal with, which are different from those of the other. Therefore, logically, the two "self's" cannot be the same "self", and so the 'copy' of the original is not the decent copy of the original's self. Kurzweil recognized the difficulty, but did not give a thought one step further. "If we copy me and then destroy the original, that's the end of me, because as we concluded above the copy is not me." [3] Having seen the difference between himself and his copy, he still called it the "copy". Since the "copy" of you is somewhat different from you, it should not be called a "copy" in the first place!

"Self" is always a singular. It is never a plural. There is at most one "self" in the world at anytime. A perfect mind-copying or reverse engineering is not possible. Pursuing immortality through "brain-copying" is thus not feasible.

REFERENCES

[1] Hall, J. Storrs. Beyond AI – Creating the conscience of the machine. Prometheus Books, New York, 2007
[2] Hofstadter, Douglas R.. Godel, Escher, Bach – An eternal golden braid, 1999, Basic Books, Inc.
[3] Kurzweil, Ray. The Singularity Is Near, When humans transcend biology. Penguin Books, New York, 2005
[4] Kurzweil, Ray. The Age of Spiritual Machines – When computers exceed human intelligence. Penguin Books, Middlesex, England, 1999
[5] Kurzweil, Raymond . The Age of Intelligent Machines. The MIT Press, Cambridge, Massachusetts, 1990
[6] Minsky, Marvin. The Society of Mind. Touchstone, Simon & Schuster, New York, 1986
[7] Minsky, Marvin. The Emotion Machine – Commonsense thinking, artificial intelligence, and the future of the human mind. Simon & Schuster Paperbacks, New York, 2006
[8] Penrose, Roger. The Emperor's New Mind. 1999, Oxford University Press, 1999
[9] Penrose, Roger. The large, the small and the human mind. Cambridge University Press, 2000
[10] Penrose, Roger. Shadows of the mind, - a search for the missing science of consciousness. Oxford University Press. 1994
[11] Richards, Jay W. (editor). Are We Spiritual Machines? – Ray Kurzweil vs. the critics of strong A.I.. Discovery Institute Press. 2002
[12] Russell, Stuart & Norvig, Peter. Artificial Intelligence – A modern approach, Prentice Hall, New Jersey. 2003
[13] Searle, John R.. The Mystery of Consciousness. The New York Review of Books, New York, NY. 1997

EMERGENT CHANGES IN IT OUTSOURCING CLIENT-VENDOR RELATIONSHIPS

Mohammed H. A. Tafti, Department of Information Technology and Quantitative Methods, Hofstra University, Hempstead, NY, (516) 463-5720, Mohammed.H.Tafti@hofstra.edu

ABSTRACT

Information technology (IT) outsourcing has considerably transformed from the way it was practiced some three decades ago. A substantial change in various attributes of clients, vendors, and their relationships have occurred during a relatively short history of IT outsourcing. For example, clients and vendors are more open to sharing knowledge and engaged in strategic partnership in performing various IT functions. The purpose of this paper is to present an analysis of the changing characteristics of client-vendor relationships in the playing field of IT outsourcing as indicated in the literature.

INTRODUCTION

Outsourcing IT has become an integral component of corporate strategy in today's organizations (Krishnamurthy *et al*, 2009). A major reason organizations outsourced their IT activities in the past was to reduce their high cost of application development (Lee, 2001). As high as 80% of the respondents in a previously conducted survey cited the desire to reduce IT costs as a key reason for outsourcing all or part of their IT operations (Lacity & Willcocks, 1998). Another common reason for resorting to IT outsourcing has been a firm's strategy to focus its resources on those activities that are considered its strengths, often referred to as core competencies. Organizations further support their outsourcing decisions by reasoning that vendors possess economies of scale that are unavailable to an individual firm. Vendors have access to more powerful equipment, are able to negotiate better deals with hardware and software providers, employ a very specialized labor force, and operate at much higher levels of production (Barthelemy, 2001).

However, considerable shifts in role and significance of these incentives and subsequent changes in various IT outsourcing attributes have created a new landscape of contemporary IT outsourcing practice (Fish and Seydel, 2006). For example, among all reasons to outsource the one that has emerged as the prime reason in recent years is to gain competitive advantage through partnership by sharing information and knowledge (Willcocks *et al*, 2004). While IT outsourcing practices during late 80's and 90's have predominantly encompassed routine, transaction-based type activities such as program coding and help desk activities, more and more of recent outsourcing contracts are on higher-level, knowledge-based IT activities such as business processes and ERP.

Besides the shift from cost-justification, there exist a host of other interesting changes in the trends of IT outsourcing. For example, as organizations gained more experience in outsourcing their IT activities and as outsourcing companies became more mature in offering IT services, new types of outsourcing deals and contracts have emerged. ***What are specific changes in characteristics of IT outsourcing client-vendor relationships?*** This "work-in-progress" paper is a preliminary attempt to shed some light on these questions by exploring relevant elements of a framework that can explain the changing landscape of IT outsourcing.

CHANGING ATTRIBUTES OF CLIENT-VENDOR RELATIONSHIPS

Several transformations in IT outsourcing attributes as related to clients, vendors, and client-vendor relationships are noticed. Although most, if not all, of factors that have influenced the decision to outsource in the 80's and 90's are still valid, there is a shift in the order of significance of these factors. Also, new attributes such as knowledge-sharing and strategic alliance are gaining significant role in shaping the IT outsourcing landscape (Lee, 2001).

Since the early era of IT outsourcing in 1980s when outsourcing clients and vendors would mostly deal at "arm's length", there have been considerable changes in client-vendor relationships. For example, while there is still a need for signing an outsourcing contract, the terms and contents of more recent contracts are noticeably different from those of early 1990s. Clients and vendors are more open to sharing knowledge and interested in joint collaboration in performing IT functions. Following presents a discussion of changes in some of the major areas of client-vendor relationship.

***Strategic Collaborative Work*:** The relationship between IT outsourcing vendors and clients is gradually shifted from simply an arm's length transaction in search of higher cost efficiency to global collaborative work (Cederlund et. al., 2007). The shift to strategic collaboration has elevated benefits of IT outsourcing beyond the limits of arms-length transaction strategy, and when implemented properly, provides both clients and vendors greater rewards than merely lower development costs (Koh, Ang, and Straub, 2004). However, proper implementation of effective collaboration is a challenging task, particularly for offshored projects, due to status differences and rigid boundaries that exist among organizations. Despite this challenge, effective collaborations have been achieved through open communication among participants attempting to freely share their viewpoints and "re-negotiate" boundaries and status differences (Levina and Vaast, 2008).

***Relationship Management*:** As the relationships between IT outsourcing clients and vendors became more sophisticated and complex, organizations required adopting a systematic approach to effectively manage their outsourcing relationships. Lessons learned from the past outsourcing deals as well as the changes in outsourcing motivation, among other factors, have led to a noticeable maturity in, and effective management of, outsourcing relationship (Solli-Sæther and Gottschalk, 2008). Traditional approach to implementation of IT outsourcing deals has emphasized preparing a complete and comprehensive formal contract in order to minimize associated risks. While this view is still relevant and valid, formal contracts are more and more being complemented by unwritten inter-organizational exchanges that are conducive to building mutual trust and social identification. Managing client and vendor relationship in order to enhance mutual trust and commitment have demonstrated positive impact on IT outsourcing success.

***Maturity of Outsourcing Industry*:** In the short history of IT outsourcing, the industry has emerged as a highly competitive and profoundly established industry with professional associations, certification programs, and strong public relations media. The industry participants can collectively influence shaping the structure, strategy, magnitude, and direction of client-vendor relationship (Grimshaw and Miozzo, 2006).

The Outsourcing Institute is one of the major professional associations established in 1993 and "tracks and forecasts the rapid evolution of outsourcing while providing new services and programs to assist buyers and sellers of outsourcing services as well as industry influencers." (http://www.outsourcing.com). The institute serves as the industry catalyst, as hub for information exchange, and as a "go-to" place for collaboration. It provides free membership, conducts research, and provides publications and information on the outsourcing industry. The IT outsourcing industry is adequately established to provide practice guidelines and offer professional certifications (Pratt, 2006). Furthermore, there are various collective efforts to strengthen the industry stands on various fronts and address industry-related issues. For example, there is a recent initiative by a service firm to launch "outsourcing bill-of-rights" covering issues ranging from transparency in the financial viability of an IT vendor in order to gain industry support in establishing standardized rules of engagement for clients and vendors in their IT outsourcing deals (Kolbasuk McGee, 2009).

***Globalization*:** Initial IT outsourcing contracts with national vendors such as IBM and EDS has gradually been complemented and/or replaced by outsourcing arrangements with vendors outside the United States. Sending the work overseas, called offshoring, has gained significant momentum by mid-90s due to the sudden rush to achieve Y2K compliance as well as the dot.com "bandwagon" phenomenon resulting in severe shortage of qualified programmers. India has become a major player in IT outsourcing due to abundance of English-speaking, talented IT professionals, particularly computer programmers (Davies, 2004). Most recently, the IT outsourcing market has become global with several major players including India, China, Russia, European and southern American countries among many others.

***Vendor's Host Country Initiatives*:** Alongside industry-wide associations and professional organizations, closer interest and support of the vendor's host country government as well as international organizations have emerged to be substantial in forging successful deals between vendors and clients. For example, among major challenges for organizations engaged in IT outsourcing are protecting privacy of the clients and information security of the firm as well as safeguarding the firm's intellectual property (Vijayan, 2004). There are steps being taken to legally, and through various regulatory agencies, establish favorable regulations and law enforcement procedures. For example, in the fall of 2006 the Indian Cabinet cleared the amendment to the "IT Act 2000" that paved the way to subsequently introduce a Bill addressing security and privacy concerns (Rustad and Koenig 2007). Also, the *World Trade Organization* has added regulatory procedures on intellectual property (*Trade-Related Aspects of Intellectual Property Rights*) that must be adhered by all member countries.

***Knowledge-Sharing*:** The early outsourcing practices involved mainly handing over routine and structured IT functions such as program coding to vendors to complete and deliver the job based on contractual terms and conditions; a typical arm's length transaction. Recent outsourcing activities, on the other hand, encompass a considerable level of knowledge-sharing between clients and vendors (Gottschalk and Solli-Saether, 2005). Knowledge-sharing strategy for IT outsourcing is more encompassing than mere exchange of information. Clients are not simply passive recipients of completed jobs; they proactively are involved in, and influence, all stages of design and development process. Furthermore, IT outsourcing has become a possible and plausible way to integrate and enhance client-vendor knowledge. This perspective would help nurturing the organizational learning capability. Therefore, the shift to knowledge-sharing strategy implies a considerable change in IT outsourcing practice in terms developing and using knowledge management skills. To take advantage of the full benefits of knowledge-sharing, however, both service provider and service receiver must be ready and willing to collaborate to build a successful partnership in creating, sharing, and using knowledge (Willcocks *et al*, 2004).

SUMMARY AND CONCLUSION

During the past three decades IT outsourcing has become an integral component of corporate strategy in today's organizations. Review of outsourcing literature indicates a number of noticeable changes in the relationships between IT outsourcing client and vendors. For the purpose of gaining a better understanding of the nature and direction of these changes, this "work-in-progress" presents a preliminary literature review exploring changes in six attributes of client-vendor relationships.

REFERENCES

1. Barthelemy, J. (2001). The Hidden Costs of IT Outsourcing, *MIT Sloan Management Review*, 42(1): 60-69.

2. Cederlund, J., Kauffman, R. J., Kohli, R., Markus, M. L., Melville, N. and Pavlou, P. A. (2007). Global Sourcing of IT Services: Necessary Evil or Blessing in Disguise? *Communications of AIS*, (2007)19: 278-293.

3. Davies, P. (2004). *What's this India Business? Offshoring, Outsourcing, and the Global Services Revolution*. Nicholas Brealey International, London.

4. Fish, K. E. & Seydel, J. (2006). Where IT Outsourcing is and Where it is Going: A Study Across Functions and Department Sizes, *Journal of Computer Information Systems*, 46(3): 96-103.

5. Gottschalk, P. & Solli-Saether, H. (2005). Critical Success Factors From IT Outsourcing Theories: An Empirical Study, *Industrial Management & Data Systems*, 105(6): 685.702.

6. Grimshaw, D. and Miozzo, M. (2006). Institutional Effects on the IT Outsourcing Market: Analyzing Clients, Suppliers and Staff Transfer in Germany and the UK. *Organization Studies* (27)9: 229-1259.

7. Koh, C., Ang, S. and Straub, D.W. (2004), “IT outsourcing success: a psychological contract perspective”, *Information Systems Research*, 15(4): 356-73.

8. Krishnamurthy, K., Jegen, D., and Brownell, B. (2009). Strategic Out-Tasking: Creating ‘‘win–win’’ outsourcing partnerships. *Information & Management,* 46(2009): 42-51.

9. Kolbasuk McGee, M. (2009). Outsourcing Bill Of Rights Proposed. *InformationWeek*, http://www.informationweek.com/news/services/ outsourcing/showArticle.jhtml?articleID=212902958&cid=nl_IWK_ daily_H

10. Lacity, M. & Willcocks, L. (1998). An Empirical Investigation of Information Technology Sourcing Practices: Lessons from Experience, *MIS Quarterly*, 22(3): 406-408.

11. Lee, J. N. (2001). The Impact of Knowledge Sharing, Organizational Capability, and Partnership Quality on IS Outsourcing Success, *Information & Management*, 38(5): 323-335.

12. Levina, N. and Vaast, E. (2008). Innovating or Doing As Told? Status Differences and Overlapping Boundaries in Offshore Collaboration. *MIS Quarterly*, (32)2: 307-332.

13. Pratt, M. K. (2006). Certified Outsourcing Professional. *Computerworld*, (40)39: 38-41.

14. Rustad, M. L. and Koenig, T. H. (2007). Negligent Entrustment Liability for Outsourced Data. *Journal of Internet Law*, (10)10: 3-6.

15. Solli-Sæther, H. and Gottschalk, P. (2008). Maturity in IT outsourcing relationships: an exploratory study of client companies. *Industrial Management & Data Systems*, (108)5: 635-649.

16. Vijayan, J. (2004) “Offshore outsourcing poses privacy perils”, *ComputerWorld*, Vol. 38. No. 8, p.10.

17. Willcocks, L., Hindle. J., Feeny, D., Lacity, M. (2004) IT and Business Process Outsourcing: The Knowledge Potential. *Information Systems Management*, (21)3: 7-15.

TECHNICAL AND USABILITY TESTING OF A MEDICAL INTERACTION WORKGROUP APPLICATION SUPPORT SYSTEM (MIWASS)

James A. Rodger*
Indiana University of Pennsylvania, MIS and Decision Sciences, Eberly College of Business & Information Technology, Indiana, PA 15705

ABSTRACT

The Medical Interaction Workgroup Application Support System (MIWASS) was run on portable handheld Personal Digital Assistant (PDA) devices. The MIWASS dental and veterinary software modules were evaluated by users from a group of subjects attending national and international conferences in the United States. The end users agreed to use the application in exchange for a chance to win a PDA mobile device. The user group consisted of 36 dentists and 36 veterinarians.

Keywords: Decision Support System, Task Technology Fit, Usability, Ease of Use

BACKGROUND

The study began in 2006 to explore the implementation of the portable platform for electronic collection of operational-level dental and veterinary information at the point-of-care. During the first year of this study, we developed the MIWASS dental and veterinary modules. These modules, based on requirements derived from Army dental and veterinary forms, show promise in confirming hypotheses that the platform can improve dental and veterinary care. These improvements would result from an increase in the accuracy of information necessary for medical situational awareness.

The MIWASS dental and veterinary modules will minimize errors in point-of-care data entry resulting from hand transcription. The use of these modules, integrated with a platform installed on a handheld device, will also increase flexibility in data collection during fieldwork. These hypotheses arise from the findings documented in previous studies and articles.

According to one such article, as advances in information technology increase the potential to collect, store, and share information, entire fields of research dedicated to the integration of technology with practice have developed. In the medical context, medical informatics has developed as a legitimate field of study, aimed at finding creative solutions to digitizing and simplifying medical information collection, storage, and presentation. The medical informatics field "deals with the acquisition, compilation, storage, and appropriate distribution of clinical or other health data" (Schomer 1999). Much of the drive to introduce information technology into the medical field stems from numerous studies that have found varying degrees of human error in medical settings, and the realization of potential inefficiencies inherent in current methods.

Though estimates of impact have varied from study to study, what is clear is that medical errors have a significant and detrimental effect on patient care. Medical mistakes were the focus of a widely-noted 1999 national report that estimated 44,000 to 98,000 Americans die each year because of errors and low-quality care (O'Leary 2006). Another study sampled pediatric immunization records in a major hospital system during a 12-month period (Wilton and Pennisi 1994). This facility used hand-written documentation, which was later transcribed on-site by a medical transcriptionist. The authors found that 22.5% of the records examined were not up-to-date. Additional examination of these cases revealed that at least 10.2% (214 records) had at least one transcription error. Moreover, 72 charts had immunization information documented in sections that were not routinely transcribed (Wilton and Pennisi 1994).

A 2003 British study provided journals to ten general medical practices for a two-week period. During this time, healthcare workers at these facilities were instructed to document any observed errors, defined as "an event that was not completed as intended and/or meant that work was disrupted in some way" (Rubin, Chinn, and Richardson 2003). Of relevance to this research is the preponderance of prescription and communication errors these respondents discovered. Approximately 42% (397) of the recorded errors related to prescriptions. Furthermore, 30% of the errors were communication errors, with missing case notes most prominent.

Medical transcription errors have been identified as the eighth leading cause of death in the United States and can lead to iatrogenic morbidity and mortality (Schomer 1999). These errors can occur at the point of service as handwriting errors, or in the data conversion phase as transcription errors. Examples of poor handwriting in a medical setting have led to the prescription of inappropriate medications and synergistic drug interactions (Rodger, Pendharkar, and Paper 1999). Exacerbating this problem is the dereliction of reporting duty by medical staff. A survey of thirteen US hospitals found that over 50% (850) of medical staff witnessed a medical mistake made by a colleague. However, less than 10% discussed the error with the colleague (Jablow 2005).

Realization of the prevalence of medical errors, or adverse events, has inspired lawmakers and researchers to propose reform. In particular, lawmakers have pushed hospitals to introduce computerized physician-order entry (CPOE) systems to aid in medication prescription entry (Davies 2005). Despite the increased interest in medical informatics, a 2005 study asserted that only five to nine percent of US hospitals have implemented some form of CPOE (Koppel, et al. 2005).

While some of this delay in transition can be attributed to variations in organizational rates of technology acquisition and implementation, the lack of widespread acceptance of CPOE and other medical point-of-care technologies underlies a greater concern. While there is a body of literature comparing the various methods of data collection and entry, the relative benefits of each approach are often unclear in some studies. Conflicting studies have left open the possibility that direct physician input of medical information may introduce greater error, relative to handwritten and transcribed records.

For example, a 2005 study found that a CPOE system introduced at the Hospital of University of Pennsylvania increased the probability of twenty-two types of "situations," or errors. These errors principally involved failure due to data and system fragmentation, and human-machine interface errors (Koppel, et al. 2005). However, this study is limited insofar as it compared the hospital's current CPOE system against a previous system, and not handwritten-transcription methods (Davies 2005).

Admittedly, a meta-evaluation by Hogan and Wagner revealed that little research had compared direct physician data entry systems against the paper standard (Hogan and Wagner 1997). These authors did, however, acknowledge that structured data forms significantly improved accuracy relative to unstructured forms. Kuhn, et al. (1992), supports this. Even fewer studies have compared the error rates of handheld point-of-care devices versus the central terminal standard. Despite the inconclusive debate, the landmark Institute of Medicine (IoM) report compels the medical field to adopt point-of-care direct physician input (Dick and Steen 1991).

Creating a point-of-care system requires the entry of information during the physician-patient interaction. Therefore, terminal-based computing does not seem to fit the IoM requirements. However, the use of PDAs offers great promise in terms of mobility, convenience, and accuracy.

The MIWASS team has developed PDA modules capable of capturing, storing, and processing data. Improved data acquisition and transmission capabilities will allow connectivity to other systems. The MIWASS dental and veterinary modules have been programmed with preformatted, structured forms. We submit that this innovation will improve data quality, efficiency, and usability.

The PDA-based modules developed for this effort will add the necessary flexibility required for fieldwork, which will be a boon to military veterinarians. In fact, the nature of fieldwork necessitates the flexibility provided by PDAs, which paper-based and laptop-based technologies lack. Pascoe, Ryan, and Morris successfully introduced a modified PDA for use in an ecology study in Kenya (2000). From this software trial, these authors summarized four characteristics of fieldworkers' work patterns that necessitate PDAs. These are dynamic use, limited attention, high-speed interaction, and context dependence. For the dental component, one can assume that dynamic use and limited attention apply.

Prior to software development, several limitations and benefits of handheld devices were considered. There is a wealth of case studies in which healthcare practitioners were provided with handheld devices. Their criticisms and positive comments are similar in each study. In Chiang, et al., ten medical residents tested a handheld device that was "hot-synched" to a patient records database (2003). Using structured forms, residents were able to enter patient information directly into the handheld device, which was then downloaded to a central workstation via a cable link. The chief criticisms among this cohort were slow CPU speed, particularly with low-end PDAs, and a "cumbersome user interface" (Chiang, et al. 2003). Another study echoed these criticisms. Respondents indicated that the screen size was too small for text-intensive purposes, and the system was unstable at times (Carroll, Saluja, and Tarczy-Hornoch 2002).

Finally, a 2001 study of 72 nurses in a Taipei hospital revealed that a statistically significant number of respondents found the PDA application easier to operate than the traditional terminal workstation (Chang, et al. 2003). However, they also cited interface concerns. This study went a step further by introducing technology acceptance questions, based on the Technology Acceptance Model (TAM) (Davis 1989). Here, the authors found that nurses were statistically less willing to accept the PDA technology than the terminal technology (Chang, et al. 2003). However, this study was limited insofar as it only provided nurses four four-hour sessions with the PDA. As Hubona and Burton-Jones indicate, there is a moderate direct effect of time since first use on beliefs about usefulness (2002).

A review of the literature on current PDA use in the medical field reveals untapped potential. To date, a majority of previous handheld device use by physicians has been limited to decision support. Organizations such as Handheldmed, Inc. (www.handheldmed.com), Lippincott Williams & Wilkins (www.lww.com/stores/pda/pdalist.html), ANMLDR (http://www.anmldr.com/), and the University of California–Davis School of Veterinary

Medicine (http://vetpda.ucdavis.edu/Projects/VetPDA.cfm) have developed applications to be run on a number of handheld platforms. However, these applications are generally limited to reference materials, dosage guides, and differential diagnostic assistance.

We believe that by integrating standardized forms into the platform, a robust, quality-driven system will replace redundant and inefficient paper data collection methods.

METHODOLOGY

We set out to demonstrate the usability and ease of use of the MIWASS dental and veterinary software modules by conducting usability testing as part of the study. This usability testing included the distribution of a survey to trained users of the MIWASS dental and veterinary modules. This survey incorporated a six-point Likert scale for measurement purposes. The survey featured 16 closed-ended questions.

Our measurement method for user evaluations is supported by the literature (Folmer and Bosch 2004; Chen, et al. 2004; Koppel, et al. 2005). In an article discussing methods for measuring usability, Folmer wrote that closed-ended surveys/questionnaires are an excellent way to measure user usability. He explained that questionnaires are widely used and reliable. Folmer lists the following as reliable questionnaires: Questionnaire for User Interface Satisfaction (QUIS), Perceived Usefulness and Ease of Use (PUEU), and User Satisfaction Instrument (USI) (Folmer and Bosch 2004). The MIWASS team utilized these questionnaires as references for survey development.

The basic rationale of the MIWASS team is that a survey/questionnaire administered to users will provide important feedback about the dental and veterinary modules developed for the PDA platform. After gathering such feedback, MIWASS software developers will consider using the survey results as the basis for improvements to the MIWASS dental and veterinary modules. This could enhance the quality, efficiency, and usability of the dental and veterinary modules.

In addition to using a survey to measure user satisfaction with the MIWASS dental and veterinary modules, the MIWASS team embedded code in the software to gain insights about the efficiency of the modules. This embedded code measured aspects such as the length of time users took to complete tasks. Hilbert provided justification for the use of embedded coding. He noted that "built-in" codes are an excellent way to provide summary statistics pertaining to user behavior (Hilbert and Redmiles 2000). We applied the results of the embedded code and survey to measure the quality, efficiency, and usability of the existing MIWASS dental and veterinary modules. Team members performed multiple-regression to assess the contribution of perceived usefulness and perceived ease of use on overall intent to use. Table 1 illustrates the results of the multiple-regression analysis.

Table 1. ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	31.922	14	2.280	16.991	.000(*)

	Residual	7.515	56	.134		
	Total	39.437	70			

* significant at 0.000 level

The results indicate that overall user evaluations of the IT application use are determined by the TTF model item's ease of use, usefulness, performance impact, and technological fit. The F value was 16.991, and the model was significant at a p < 0.001 level of significance. Table 2 shows the R^2 for the model was 0.809, which indicates that model independent variables explain about 81% of variance in the dependent variable.

Table 2. Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.900	.809	.762	.366	1.727

In addition to the model summary information, Table 3 shows a table of coefficients for each of the independent variables (items not measuring Intent to Use), along with t-test, standard error, and significance. Questions 1, 4, 5, 7, 8, 11, 14, and 16 were significant at the 0.05 level of significance.

Table 3. Coefficients for Each of the Independent Variables

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.552	.480		-1.151	.255
	Q1	.271	.103	.320	2.639	*.011
	Q2	-.190	.115	-.204	-1.651	.104
	Q3	.096	.071	.107	1.347	.184
	Q4	.232	.114	.243	2.042	*.046

	Q5	-.246	.087	-.306	-2.823	*.007
	Q6	.083	.096	.098	.861	.393
	Q7	-.191	.086	-.228	-2.225	*.030
	Q8	.244	.107	.232	2.278	*.027
	Q9	.081	.088	.083	.917	.363
	Q11	.220	.091	.215	2.424	*.019
	Q12	-.025	.046	-.041	-.547	.587
	Q14	.240	.114	.220	2.116	*.039
	Q15	.014	.122	.012	.116	.908
	Q16	.280	.100	.295	2.794	*.007

* 0.05 level of significance

CONCLUSIONS

The contribution of our research can be summarized as follows: we developed an instrument to measure the usefulness of the MIWASS dental and veterinary modules in healthcare, and we successfully tested the TTF model on those modules. The TTF model is a continuously evolving model for measuring user evaluations of information systems. Though earlier studies (Goodhue and Thompson 1995; Goodhue 1995) focused on developing a multi-dimensional construct for user evaluations of TTF, very little information was provided on how to measure technology, task, and individual characteristics constructs. We used the TTF model to develop an instrument for obtaining user evaluations of the MIWASS dental and veterinary modules. We have shown that the major factors that impacted the overall use of the modules were that they were easy to use, useful, and fit users' technological needs. Our results indicate that TTF, along with individual characteristics, has an impact on user evaluation of healthcare IT.

Survey participants were quickly and successfully trained to use the MIWASS dental and veterinary modules for and were satisfied the modules performed most of their data collection functions very well. Participants also indicated the modules could be useful tools in collecting and disseminating data and had features that would allow users to obtain, evaluate, and present information more efficiently than previous methods. Overall, participants indicated the MIWASS dental and veterinary modules had significant potential digital data collection utility. Based on these results, we recommend that the MIWASS dental and veterinary modules be introduced to the workforce to enable the digitization of dental and veterinary records.

References: Available Upon Request

TOWARDS AN INTEGRATED FRAMEWORK OF WIKIPEDIA ADOPTION : MOTIVATIONAL FACTORS OF CONTRIBUTORS AND UTILIZERS

April Bockmann, Richard Stockton College of NJ, bockmana@go.stockton.edu
Aakash Taneja, Richard Stockton College of NJ, aakash.taneja@stockton.edu

ABSTRACT

Wikipedia is a comprehensive, frequently used source of information to which anyone can contribute articles that anyone else can edit. While its concept has been embraced by many, skeptics have challenged its quality and reliability. The popularity of Wikipedia has opened new challenges and avenues for researchers. In this study, we derive an integrated model of the antecedents of individuals' involvement with Wikipedia and the drivers of Wikipedia's success.

INTRODUCTION

In today's world, there is no excuse for not being able to find a desired piece of information. The World Wide Web gives access to limitless information, and conceivably contains every answer to any question. One website that can be used to sift through and organize this seemingly endless data is Wikipedia.org. It consists of wikis which are edited, deleted, or modified by the public and then published on the website for the whole community to share. Wikipedia is a rather radical idea, compared to its conventional counterparts. While its concept has been embraced by many, skeptics have challenged its quality and reliability. These opposing viewpoints have raised new challenges to adequately explore this phenomenon.

The objective of this research is to compliment and further develop the understanding of individuals' involvement in online communities like Wikipedia. Although past research has utilized various theoretical perspectives in isolation to study this phenomenon, there is a need to integrate these fragmented pieces into a unified framework that pertains to both producers and consumers of Wikipedia. Specifically, this research seeks answers to the following questions: i) What are the drivers behind Wikipedia's success? ii) What are the antecedents of knowledge contribution by individuals in Wikipedia? iii) What are the antecedents of knowledge utilization by individuals from Wikipedia?

The rest of the paper is organized as follows. The next section presents an overview and purpose of social computing. It is followed by a discussion of "What", "Who", "How" and "Why" of Wikipedia to develop the research model and hypotheses. Lastly we present our conclusions.

Social Computing and Wikipedia

Social computing technologies are a set of "applications and services that facilitate collective action and social interaction online with rich exchange of multimedia information and evolution of aggregate knowledge" (Parameswaran and Whinston 2007). Some of the well known social computing systems are Wikipedia, YouTube, MySpace, and FaceBook. Of particular importance is Wikipedia, an online encyclopedia that facilitates collective action to provide information on a wide variety of topics and ideas. There is no denying the force that Wikipedia has presented in the online community. Its convenience accounts for its popularity for knowledge seekers, and ease of use along with sense of purpose motivate millions of contributors. In order to understand the success of Wikipedia we have to consider “What” Wikipedia is setting out to do, “Who” are its stakeholders, “How” does it work and finally “WHY” people are involved with the project.

The What?

Wikipedia is a comprehensive, frequently used source of information for millions of users to which anyone can contribute articles that anyone else can edit. Wikipedia appeared as an outcome of Nupedia which aimed to produce an encyclopedia by involving highly qualified contributors and a strict peer review process. However, the slow pace of writing articles led to abandonment of Nupedia. This weakness of Nupedia (slow pace of writing) became the strength of Wikipedia, and its strength (peer review process) turned out to become Wikipedia’s weakness.

The Who?

An important step in evaluating the technology is to understand the psyche of the average Wikipedian. As observed by Hara and Shachaf (2009), the participatory nature of the social Web has given birth to communities where users often become “prosumers,” who are consumers and producers at the same time. However, it is pertinent to mention that a consumer may or may not be a producer. This acceptance of free-riders adds to the dimension of Wikipedia.

How?

Wikipedia gets many people to contribute bits of information and then pieces them all together to form cohesive articles. The fates of contributions are determined by verifiability, and the self cleaning nature of the project works in its favor. Since there are many people willing to dedicate time to strengthen and correct articles, there are more opportunities to find mistakes. One thing in Wikipedia's favor is that it is constantly updated and one can get information almost in “real time” (Thompson 2005). Also, since all of the history of an article is stored and viewable by the public, users can get multiple perspectives on a topic which illustrate the train of thought of a subject, and is helpful in forming personal opinions.

Why? – The Motivation

Members’ motivation to actively get involved in the community in order to contribute, or to consume knowledge is a key ingredient of the community’s success. In addition to technological characteristics, drivers of individual’s involvement with Wikipedia can be evaluated in terms of both intrinsic and extrinsic motivations. The two stakeholders involved with Wikipedia (producers/consumers) have a different set of desires behind their adoption and use of Wikipedia, and the success of Wikipedia primarily depends on the continued involvement of “both” of them.

PRODUCERS / CONTRIBUTORS

The key dependent construct for contributors is related to knowledge contribution, defined as the willingness of individuals to voluntarily contribute in Wikipedia. It is hypothesized to be influenced by intrinsic and extrinsic motivation, technological characteristics and factors inhibiting its usage as discussed in the following sections.

Intrinsic Motivation

Community identification: It is a form of social identification in which individuals identify themselves as a member of a community (Ashforth and Mael 1989) and takes the community's interest to heart (Law and Chang 2008). According to social identity theory, individuals are more likely to support a community once they identify themselves with it (Wang, Lin et al. 2008). In the case of Wikipedia, if the individuals have a sense of belongingness with Wikipedia, they are more likely to post / edit entries in it. Therefore,

H1: Community identification is positively related to knowledge contribution in Wikipedia.

Altruism: It reflects an individual's goal to increase the welfare of others. Altruistic individuals help others because they receive intrinsic benefits and feel good when doing so as compared to receiving any external rewards. Many contributions to Wikipedia can be considered altruistic acts as Wikipedians invest their time and effort for the satisfaction of giving, and without any monetary compensation (Kuznetsov 2006). Therefore,

H2: Altruism is positively related to knowledge contribution in Wikipedia.

Cooperation Norm: It represents a degree of consensus in the social system and refers to a general spirit of collegiality and mutual understanding among the individuals. In terms of Wikipedia, often known as wikilove, it includes respecting other contributors, encouraging newcomers, following Wikipedia policies etc. Individuals may be more willing to contribute when adherence to the philosophy of wikilove are prevalent in the community. Therefore,

H3: Cooperation Norm is positively related to knowledge contribution in Wikipedia.

Commitment: It reflects the desire to help other members of the community. As observed by Nov and Ye (Nov and Ye 2008), commitment to the community has been shown to motivate individuals to share knowledge in various settings. Wikipedians have shown commitment to the project by forming a system to allow everyone to do their share. They put forth effort to strengthen weak articles and to combat ill-intentioned behavior. Therefore,

H4: Perceived commitment is positively related to knowledge contribution in Wikipedia.

Opportunistic behavior: It is defined as an individual's consideration of self – interest without recognizing the consequences to others. In terms of Wikipedia, it involves parties altering their wiki's to enhance their image, or removing content that paints them in a negative light. The prevalence of opportunistic behavior tends to undermine the work of Wikipedians, leading to their reduced level of participation in the community. Therefore,

H5: Opportunistic behavior is negatively related to knowledge contribution in Wikipedia.

Reputation: Developing reputations and online identities within the community is a major contributing factor for quality edits to the website. Wikipedians work to gain trust and respect from their fellow contributors (Kuznetsov 2006) and their desire to be valued by the rest of the community motivates them to contribute quality edits. Therefore,

H6: Need for reputation is positively related to knowledge contribution in Wikipedia.

Extrinsic Motivation

Reciprocity: It reflects the individuals' wish to pay back the benefits received or their expectation in terms of receiving future benefits from the community to which they contribute. As Wikipedia depends on its members for success, in order for Wikipedians to reap the benefits, they must rely on each other to achieve their goals (Kuznetsov 2006). Therefore,

H7: Reciprocity is positively related to knowledge contribution in Wikipedia.

Self-development**:** It involves learning from others in the field, receiving feedback, and enhancing one's abilities and skills. It is associated with knowledge sharing (Von Krogh, Spaeth et al. 2003) and is one of the motivations for sharing in open content project communities such as Wikipedia (e.g. (Oreg and Nov 2008)). Wikipedians who believe that they will gain useful feedback on their posts will be more inclined to contribute to Wikipedia. Therefore,

H8: Self development is positively related to knowledge contribution in Wikipedia.

Inhibiting Factors

Anti-Elitism**:** This reflects the common attitude within the committee to not blindly honor the contributions from experts. Since there is no central authority for experts to appeal to, it is often one word against the word of the group. If their work is constantly deleted, experts will become frustrated and will cease volunteering their time (Fallis 2008). Therefore,

H9: Anti-Elitism is negatively related to knowledge contribution in Wikipedia.

Exclusiveness manifestation**:** It is defined as the tendency of a group to take control by neglecting the voice of others. In Wikipedia, many times this bias develops among a small group of core editors. Sanger (2004) describes it as "a poisonous social and political atmosphere." Because of this, contributors complain of cliquishness and leave Wikipedia. Therefore,

H10: Exclusiveness manifestation is negatively related to knowledge contribution.

Technological Characteristic

Trust: Trust plays a key role for effective collaboration and plays a critical factor in developing various relationships. It reflects individuals' perception of the credibility of Wikipedia as a source for reliable information. If individuals believe the project to be worthwhile and trustworthy, they will be more inclined to contribute. Therefore,

H12: Trust is positively related to knowledge contribution in Wikipedia.

Perceived critical mass**:** Many people tend to join and participate in a community due to network effect. According to the theory of Critical Mass (Oliver, Marwell et al. 1985), people tend to perform collective actions as the number of participants increases. As observed by Wang and colleagues (Wang, Lin et al. 2008), in the context of Wikipedia, perceived critical mass will positively accelerate the willingness of users to join and contribute to this community. Therefore,

H13: Critical mass is positively related to knowledge contribution in Wikipedia.

Effort Expectancy**:** Effort expectancy is defined as the degree of ease associated with the use of the system. If the contributors find it difficult to participate in Wikipedia due to technical limitations, they'll be reluctant to contribute. Therefore,

H14: Effort expectancy is positively related to knowledge contribution in Wikipedia.

Perceived behavioral control**:** Perceived behavioral control is defined as an individual's perception of the presence or absence of factors that may facilitate or deter performance of the behavior (Ajzen 1991). If the individuals perceive themselves to possess control in terms of contributing in Wikipedia, they are more likely to contribute in Wikipedia. Therefore,

H15: Perceived behavioral control is positively related to knowledge contribution.

Regulatory control**:** Presence of appropriate rules and regulations are found to minimize different types of undesirable behaviors. With respect to Wikipedia, presence of trolls, exclusiveness manifestation, and anti-elitism are generally seen as byproducts of the lack of regulatory mechanisms. Therefore,

H16: Perceived presence of regulatory control is positively related to knowledge contribution.

CONSUMERS/UTILIZERS

The key dependent construct for consumers is related to knowledge utilization, defined as the willingness of individuals to seek knowledge from Wikipedia. Its main drivers include extrinsic motivation, environmental and technological characteristics of Wikipedia.

Extrinsic Motivation

Performance expectancy**:** Performance expectancy is defined as the degree to which an individual believes that using the system will help him or her to attain gains in performance. If users do not have Wikipedia as a tool for seeking information, they will need to search on the vast Web where information is not reviewed by as many people and is checked much less often. The content in Wikipedia is up to date and constantly being checked for quality. Thus,

H17: Performance expectancy is positively related to knowledge utilization from Wikipedia.

Knowledge needs**:** This involves meeting the needs of the consumer in terms of matching their information-seeking goals. Perceived knowledge needs will influence the individuals' intention to use Wikipedia and consist of the following facets: i) Specific Queries, ii) Current Events, iii) Recreation, and iv) Introductory Research.

H18: Perceived knowledge needs are positively related to knowledge utilization.

Environmental Characteristic

Social Influence**:** It is defined as a process of getting persuaded by important others' in an individual's environment. If these influential others accept (or reject) the use of Wikipedia, there is a high chance that the individual will adhere to their recommendations. The role of social influence is well established in various fields which includes psychology, economics, and information systems, to name a few. Therefore,

H19: Perceived favorable social influence is positively related to knowledge utilization.

Technological Characteristic

Size**:** Consumers often get attracted to Wikipedia because it contains articles on almost every conceivable topic. If the users believe that Wikipedia lacks their needed content, it will adversely influence their intention to use Wikipedia. Therefore,

H20: Perceived size is positively related to knowledge utilization from Wikipedia.

Effort Expectancy**:** As stated earlier, effort expectancy will play a role in determining an individual's intention to use any information system. If users consider Wikipedia to be hassle free, they will be more inclined to its use. Therefore,

H21: Effort expectancy is positively related to knowledge utilization from Wikipedia.

Perceived behavioral control**:** As stated earlier, if the individuals perceive themselves to possess control in terms of using Wikipedia, they are more likely to use it. Therefore,

H22: Perceived behavioral control is positively related to knowledge utilization.

Perceived information quality**:** Information quality is an important factor affecting the use of information systems, and consists of the following facets: i) Accuracy, ii) Currency, iii) Comprehensiveness, and iv) Reliability. Users will be apt to use Wikipedia if they will find the available information to be of desired quality. Therefore,

H23: Perceived information quality is positively related to knowledge utilization.

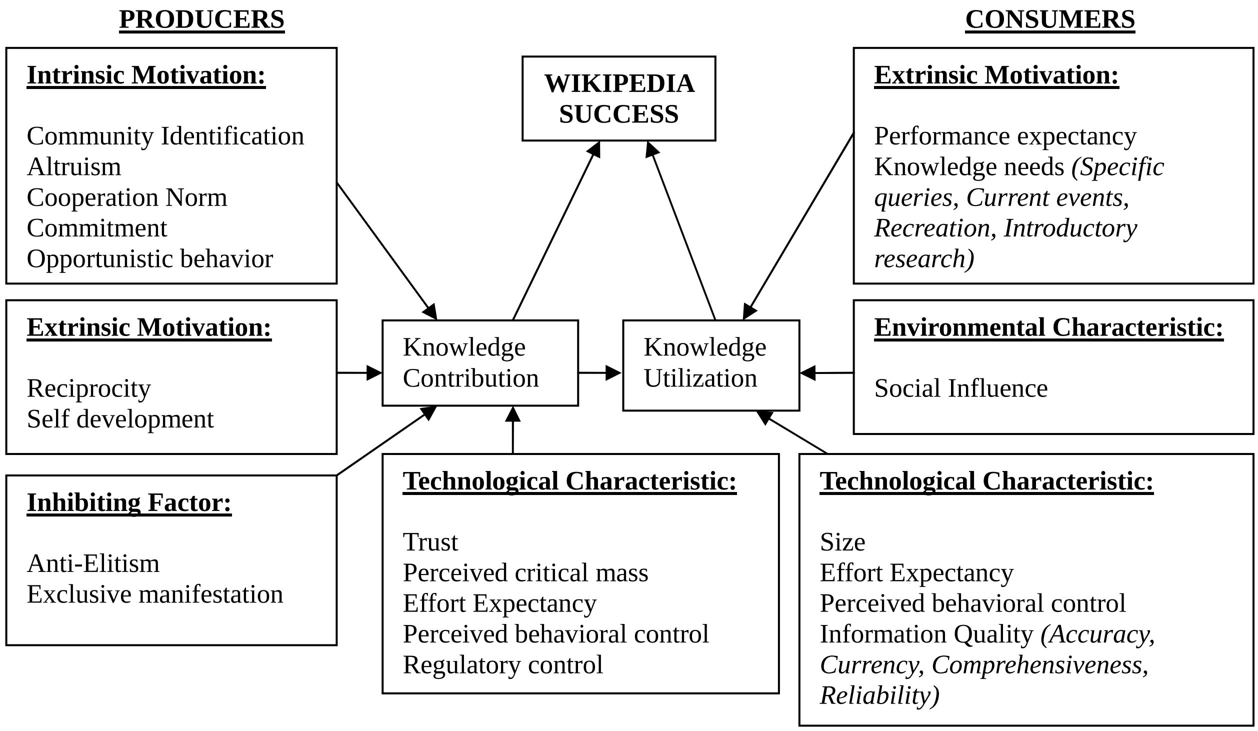


Figure 1: Research Model

CONCLUSION

Use of Wikipedia has become a de-facto standard for both students and adults. Undoubtedly, blind adoption of Wikipedia will cause undesirable effects. Still, if used properly, Wikipedia is believed to be a positive development that brings us many benefits. As the future of technology comes up on us more rapidly each day, Wikipedia will too conform to these advancements. The next wave of online information sharing is WEB 3.0, or the Semantic Web. Since the Semantic Web will directly connect data, many of the current problems that will be solved include content consistency, ease of filtering data, and ease of exchanging and reusing knowledge (Fallis 2008).

In this study, we propose an integrated model of the drivers of Wikipedia success based on the prior research. By proposing an integrated framework of adoption of an open source collaborative content mechanism like Wikipedia, we expect to make contributions to both researchers and practitioners. From an academic perspective, this study advances the stream of research on social computing and Wikipedia by unifying the viewpoints of producers and consumers of Wikipedia. By simultaneously analyzing the antecedents of intention to participate in Wikipedia in a single model, we expect to gain a richer understanding of the complex motivational issues concerning adoption of Wikipedia. Given the increasing popularity of Wikipedia, an understanding to these factors is also crucial for the practitioners' intent on exploiting this phenomenon.

References

References available upon request from Aakash Taneja

The Use Narrative Fiction in Information Technology Courses

Alexander Pelaez – Adjunct Professor of Information Technology, Hofstra University

Abstract

Review of research demonstrates the effectiveness of the use of narrative fiction to teach concepts such as Ethics (Marquis, 2003) and the study of organizations (Phillips, 1995). Researchers have argued that narrative fictions serves as a compliment to traditional text based and case based study by serving to fill in the gaps though the implicit connection readers of narrative fiction have with characters and the environment which is being created (Marquis, 2003; Shepard et al. 1997). No research had been uncovered specifically regarding the use narrative fiction for Information Technology. This paper specifically discusses the use of "Ammunition: Essentials of Leadership" in Information Technology classes. The story represents a unique opportunity for Information Technology students due to the environment and characters being entirely based in IT. The story demonstrates concepts such as groupthink, decision making, micromanagement and The OODA Decision Cycle, among other leadership concepts. Characters face challenges such as Project Deadlines, IT Strategy, IT Organizational structure coupled with technological issues such as web design, ERP systems, systems analysis, and the Software Development Lifecycle. The goal of the book was to provide students with the key elements of leadership through the challenging scenarios presented in Information Technology.

Introduction

The use of narrative fiction has played a more active role in formal business education, especially those schools that use the case method as a teaching tool (Alvarez, Merchan 1992). Researchers have argued that narrative fictions serves as a compliment to traditional text based and case based study by serving to fill in the gaps though the implicit connection readers of narrative fiction have with characters and the environment created by the story (Marquis, 2003; Shepard et al. 1997). It provides an environment by which an understanding of organizational issues can allow a practitioner to more easily relate. If narrative fiction can prove useful for business disciplines such as Ethics and Organizational Behavior, would the same hold true for Information Technology? This paper will examine the use of a specific narrative fiction, *Ammunition: Essentials of Leadership*, for use within Information Technology classes.

Use of Narrative Fiction

The use of narrative fiction in teaching business concepts such as Economics (O'Donnell 1995), Organizational Behavior (Phillips 1995; Sliwa Cairns 2007; Patient et al 2003) and Ethics (Marquis 2003) has gained in acceptance over the past few decades. The common theme of the research holds that narrative fiction compliments the traditional teaching methods.

Phillips (1995) provides a useful topology for demonstrating the complementary nature of narrative fiction in the study of organizational analysis. By creating a 2x2 grid with Non Narrative and Narrative on the horizontal axis and Non Fiction and Fiction on the vertical axis, he describes the type of approaches common for organizational analysis.

Quadrant I, Fiction and Non-Narrative, represents bodies of work such as Theories, Mathematical Models and Typologies, Quadrant II, Non-Fiction and Non-Narrative, represents work such as Survey Data, Laboratory Data, and Content

Analysis. Quadrant III, Narrative and Non-Fiction, represents work such as Biographies, Case Studies and Ethnographies. Phillips argues that Narrative Fiction, which would be represented in the fourth quadrant, is under represented in approaches to organizational analysis. The research shows that each of the quadrants represents a different method of reaching similar conclusions or providing an analysis of the hypothesis it presents. Narrative Non-fiction should therefore be represented as an equal social construction (Phillips 1995). It is "an attempt to model the world…and weave a pattern of truth in a way quite parallel to the theories of organizational analysis." (Phillips 1995) Narrative Fictions' ability to create a more complex and imaginative environment replaces simple general systems of organizational thought (Phillips 1995). This replacement creates more value to the reader thereby potentially increasing the ability to absorb the concepts being taught.

Managerial decision making represents a complex process by which managers must take into account a variety of inputs in order to solve problems expeditiously with as much accuracy as possible. This complex process requires practitioners to compliment standard and more formalized management theories with more informal sets of knowledge including personal experiences, anecdotes and stories (Alvarez and Merchan 2003). The variety of environments, places and times provide a practitioner with a unique set of experiences in which the reader can provide interpretation and use imagination within the context of management theory (Czarniawska-Joerges and Guillet de Monthough 1994). By capturing the complex nature of human relationships and psychological effects, a narrative fiction can more readily convey emotions and consequences than traditional textual based learning (Patient et.al 2003). These emotions and consequences are magnified, in some cases, based on the reader's personal experiences and perspective.

Information Technology courses are generally technical in nature; however students graduating school will find an increasing number of non-technical scenarios, such as project management, business analysis or strategy. More and more Information Technology leaders, especially CIOs are coming from non-technical backgrounds (Noda, 2008) although some argue CIOs should posses a dual background (i.e. knowledge of the business and technology) (Curran 2009).

While it is easier to see the use of literature to study the complex human relationships in areas such as Organizational Behavior and Ethics, an argument might be made against its use with disciplines such as economics or mathematics, in which Information Technology could be grouped. However, there is evidence for the success of the use of narrative fiction, within the aforementioned disciplines to generate interest in the subjects as well as demonstrate the value of theories.

The use of novels to teach economics was first introduced in 1832 by Harriet Martineau, who believed that simple tales demonstrating principles of basic economic theory would be more easily understood by a larger population (O Donnell 1989). The use of short narrative fiction enabled her to explore a single complex economic concept, providing a more comprehensive explanation in the form of a practical application as exhibited by the fictional account. These ideas were valid by the mere fact that her tales sold over 10,000 copies of the monthly series by 1834 (O'Donnell 1989). O'Donnell further shows that this use of economic novels continues.

Another analysis shows that narrative fiction is widely used to demonstrate mathematical concepts, since the popularity of science fiction ignited the imagination of rocket scientists and engineers prior to World War II. A review on Wikipedia listed 22 books in the non-fictional category (this number is only meant to demonstrate the availability of fiction books within the mathematic genre, not to provide a definitive number or validate its content).

One of the most popular forms of narrative fiction is television. The popular weekly television series “NUMB3RS”, showcases a mathematics professor, Charlie Epps, who uses complex mathematics to solve crimes with his brother Don Epps, an FBI agent. The series uses different disciplines of mathematics each episode and provides a more layman’s definition of complex equations and theories. Its uniqueness lies in the practical manner in which the explanation are given, and its use of well recognized metaphors, a key element to teaching concepts. For the first four seasons, it was the most popular show on its airing time (“Numbers”).

Its effect has been felt through the mathematical community. The Cornell Mathematics Department created a site dedicated to understanding the mathematics used on the weekly show (http://www.math.cornell.edu/~numb3rs/). Texas Instruments partnered with CBS Paramount, in association with the National Council of Teachers of Mathematics, to create “WeUseMathEveryday.com”, whose mission is “to inspire students to achieve more in math by showing how the subject is relevant to their lives”. The site conceived in 2005, after season one of the series, received 43,000 supporters who signed up for teachers kits.

Thus, there is demonstrated evidence that supports the role of narrative fiction in teaching concepts for a variety of subjects, as well as generating interest in the discipline for which the narrative is based. Furthermore, the research suggests that application of narrative fiction can be successful in complimenting traditional text based and case base studies.

Narrative Fiction for Information Technology

There exists no evidence of a purely Information Technology based narrative fiction for the purpose of teaching Information Technology Management or Strategy. While fiction has been used to create technological suspense mysteries (e.g. “Untraceable’) or science fiction (e.g. “Contact” and “Star Trek”), students of Information Technology may be deprived of a narrative fiction that demonstrates the complexities and challenges that are experienced by Information Technology managers. It could be argued that the challenges faced by IT professionals are no different than those faced in Operations, Logistics, Marketing or Finance; however, as demonstrated above, the ability of a reader to relate a story to personal experiences enhances the quality (Cohen and March 1986). Therefore, a significant benefit could be achieved by using a narrative fiction with significant Information Technology experiences.

Ammunition: Essentials of Leadership, a novel written by Alexander Pelaez, is designed to teach the concepts of leadership and strategy within an Information Technology setting. The main character Ryan starts as a lead developer who struggles with the concept of leadership as his career progresses. Ryan faces many challenges from superiors, peers and subordinates while managing his team of developers.

It is these challenges that provide an experience based understanding of the management concepts taught in traditional texts. In the first setting the character learns two key important lessons: the power of goal setting and the value of integrity.

In chapter two, the main character struggles with inheriting a project, the implementation of an inventory system, which has no chance of success. The system, acquired by the business unit, without the consent of Information Technology department, is loaded with problems. The problem is compounded by the hubris of the Information Technology department, including Ryan, believing they could get the software running in the prescribed time. It is at this point where Ryan is exposed to the concept of groupthink, its effect on decision making, and the consequences of a faulty decision making process.

As the story progresses through the next few chapters, the reader is exposed to Ryan's challenges with organizing his team and the effect of higher organizational changes on his team. Ryan works through the challenges, empowers his team, and attempts to keep projects on schedule, but finds that he is unable to overcome strong political currents for which he wasn't prepared to handle. Ryan also finds that running a technology team becomes less about managing and implementing technology, and more about dealing with the personal issues of his team and the political nature of the organization. After repeated confrontations with a peer, his peer verbally attacks a key female member of his team, Stacy, who Ryan has been cultivating as a leader. This harassment becomes Ryan's biggest challenge to this point. He must decide whether to stand up to the harassment and risk his political standing with his superiors by reporting the incident to Human Resources or concealing the incident.

Ryan, understanding his ethical and moral obligation, works with the Stacy through HR to the disdain of his superiors. Ryan, ready to accept the consequences of his decision, is stunned, when, out of retribution, another member of his team was singled out for termination, in an effort to force Ryan to conform.

Ryan eventually leaves the company, and, after a very difficult six months for Stacy, brings her along. At his new position, he is exposed to a very entrenched team focused on older technologies, and unwilling to adapt. Ryan works with the team in an attempt to motivate them; however, he soon learns that his efforts are being sabotaged by a key member of his team, Bill. Ryan, understanding how important it was to establish his position as the leader of the group in an effective yet non-combative manner, begins to restructure the group in an effort to minimize the belligerent actions of Bill. The increasing resistance by Bill eventually comes to an apex when Bill, who is the brother-in-law of the CEO, a fact not revealed to Ryan, circumvents the chain of command. Bill's efforts to circumvent Ryan are unsuccessful and he is eventually asked to resign by the CEO.

Ryan brings the team closer by introducing key elements of organizational structure: Trust, Mission, Intuition and Agreement, which are adapted from Col. John Boyd (USAF). These concepts were originally part of a work called *Patterns of Conflict*, part of a Defense Department Briefing on the effectiveness of organizations. The original terms were based on the German military Blitzkrieg strategy but were adapted for the story using English definitions. Ryan also introduces the concept of the OODA

Decision Cycle (Observe-Orient-Decide-Act) to the team, also a concept coined by Col. Boyd.

The story shows Ryan and his team implementing these concepts almost immediately during a new rollout of an inventory system, website and point of sale system. Originally, designed as three separate systems, the team would integrate the features and functions to ensure reuse of code and better integration between the two systems.

A warehouse fire destroys the inventory for the company, and bringing the old system online with updated inventory levels and shipping instructions would take weeks and incur a significant amount of manual labor. The team, understanding the concepts of OODA, quickly scrambles to get the new inventory system to production, which had many of the features required for such an emergency, but were never requested by the business.

The team would successfully implement the system and despite an attempt by the company's competition to take advantage of the situation, sales increased because of their ability to predict inventory levels, provide real time information to the customers on the website and deliver products faster. These were features the competition could not offer.

Ryan and his team would learn the importance of understanding the business and being partners with business users, as opposed to being adversaries. It is this relationship, along with Ryan's constant empowerment that makes him and his team successful, leading to Ryan's understanding of a management paradox: "The more control you try and exert the less control you actually have." (Pelaez, 2009).

The narrative fiction demonstrates the type of challenges Information Technology professionals. It serves to teach the different dynamics of the internal workings of an Information Technology group, and the conflict that occurs within technology groups and with the business units.

Current Application

Currently, *Ammunition* is used as a complimentary text for an Introduction to Ecommerce course at Hofstra University. Students have responded positively to the story line and the ensuing discussions support the premise that the narrative sparks interest in certain topics. While not an appropriate sample, 8 out of 12 undergraduate students in the class thought the book was a good compliment to the text, but more importantly, they considered it a break from the routine lessons of the class.

This particular narrative fiction would probably be better served at a graduate level, since, generally, graduate students (particularly EMBA students) have a collection of personal work experiences, which may mean better relation to the experiences in the story. This is somewhat supported by anecdotal comments made by Information Technology professionals who have read the story, who state specifically, that the experiences in the story are experiences that have personally faced or witnessed.

Limitations and Directions for Future Research

The primary limitation of this research is that there is no empirical evidence regarding the use of this particular story as a teaching tool. The story's novelty, being recently published in September 2009, poses an obvious challenge for the development of any practical statistical model; however, over the course of the next few university semesters a more comprehensive research paper could be developed. The figures presented in the above research were also non scientific, and were only meant to provide a direction of understanding and not serve as a replacement of a more formalized approach.

A further limitation stems from the inability to find different narrative fiction with an Information Technology centric environment, thus no comparison could be made between the narrative fiction presented above and some other body of work. Further research into the use of other narrative fiction might shed additional insight into the effectiveness of narrative fiction in Information technology classes and their impact on students learning.

Conclusion

This paper demonstrates the use of narrative fiction in a wide range of disciplines including Ethics, Organizational Behavior, Economics and Mathematics. It was evident from the research that narrative fiction serves as a strong compliment to traditional texts in filling the void left by traditional texts and case based approaches (Phillips 1995). The effectiveness of narrative fiction in creating a world that is more true to the world for the individuals who live in similar worlds (Phillips 1995).

The use of *Ammunition: Essentials of Leadership* provides Information Technology professionals with a collection of related experiences and therefore the lessons of leadership and management within Information Technology might be more readily absorbed.

References

Alvarez, J. L. & Merchan, C. (1992), 'The Role of Narrative Fiction in the Development of Imagination for Action', *Internation Studies of Management and Organization* **22(3)**, 27-45.

Curran, C. (2009, October 04), 'CIO Background Check: Is IT Experience Mandatory?', *CIO Magazine*.

Czarniawska-Joerges, B. & de Monthoux, P. G. (1994), *Good Novels, Better Management: Reading Organizational Realities in Fiction*, Harwood Academic Publishers.

Greenwood, M. R. (2000), 'The Study of Business Ethics: A Case for Dr. Seuss', *Business Ethics: A European Review* **9(3)**, 155-162.

Marquis, M. (2003), 'Using Literature to Teach Ethics', *Kentucky Journal of Economics and Business* **22**, 53-60.

Noda, T. S. (2008, Aug 12), 'More non-IT pros becoming CIOs', *ComputerWorld Hong Kong*.

O'Donnell, M. G. (1989), 'A Historical Note on the Use of Fiction to Teach Principles of Economics', *Journal of Economic Education* **3**, 314-320.

Patient, D.; Lawrence, T. B. & Maitlis, S. (2003), 'Understanding Workplace Envy Through Narrative Fiction', *Organization Studies* **24(7)**, 1015-1044.

Pelaez, A. (2009), *Ammunition: Essentials of Leadership*, Geolatin Publishing, Inc..

Phillips, N. (1995), 'Telling Organizational Tales: On the Role of Narrative Fiction in the Study of Organizations', *Organization Studies* **16(4)**, 625-649.

Shepard, J. M.; Goldsby, M. & Gerde, V. W. (1997), 'Teaching Business Ethics Through Literature', *Teaching Business Ethics* **1 (1)**, 33-51.

Sliwa, M. & Cairns, G. (2007), 'The Novel as a Vehicle for Organizational Inquiry; Engaging with the Complexity of Social and Organizational Commitment', *Theory and Politics in Organization* **7(2)**, 309-325.

"Numb3rs." *Wikipedia, The Free Encyclopedia*. 1 Nov 2009 Web. 3 Nov 2009 <http://en.wikipedia.org/w/index.php?title=Numb3rs&oldid=325700779>.

"'We All Use Math Every Day' Education Initiative Wins National iParenting Media Award for Best Educational Web Site" *We all Use Math Every Day,* Texas Instruments, Inc, n.d. Web. 13 Nov 2009 <http://focus.ti.com/pr/docs/preldetail.tsp?sectionId=594&prelId=et070034>

Development of Visual Display and Navigation for Small Screens based on User Behaviour and Personalisation.

Abstract

The intention of this paper is to present an application that is being developed for small screen devices in which a user can personalise content to suit their interests and learning preference. The Talking Walls® is an adaptable multimedia template application for heritage sites, museums or historical buildings. Users are able to visually explore and navigate where they are visiting via a mobile phone / handheld device, or a kiosk / DVD and online, learning about the history of a specific place / space / object, and how this has changed over time. This research explores the behavioral factors affecting how children and adults want to use hardware and software to navigate their way around content and the space. The KubeMatrix™ holds the information and material together as a content tool. This is a unique concept; it acts as a navigational tool, representing time and space, content and links in a 4-dimensional cube matrix structure, as the interface navigating through the content and enables the user to choose age range and content. Users are able to track their progress, with the KubeMatrix™ indicating the "areas" visited and the routes taken. Further research is planned to validate user behavior and attitudinal factors.

Keywords:

user-generated content; mobile devices; navigation; multimedia; heritage.

Introduction

The increasing competitiveness of the global tourism industry demands that heritage sites / museums continually seek to add value to attract visitors and the handheld / PDA tour would appear to be a major key in providing this added value. With examples such as the handheld multimedia tour at the Tate Modern and the Natural History Museum (BARRY 2005-6) consumer use of handheld devices is increasing. The development of content for these devices is an area that is of increasing importance to both the designer/creator and the consumer. It is not enough to re-purpose content already provided in other media forms, the screen size dictates new form and new content to be designed specifically for the small screen (Moll 2005).

Children are also very important in the way the content is developed, they have/are growing up with not just computers but mobile devices and social networking, the anytime, anyplace, anywhere 'Martini' generation (Ranchhod 2007). They will be the market segment that is likely to know exactly what they would like on their mobile devices, how they would like to learn and what is likely to keep them using the devices for more than just communication. Although, according to cell phone tracking company M:Metrics (PR.Newswire 2007) statistics show that 77% of people aged 55-64 have mobile phones. Where it might have been assumed that this technology would only appeal to the younger generation, these figures show that this may not be the only market to design for. The 55–64 age range may have more in the way of a disposable income, combined with possibly less responsibilities and more time to explore their cultural heritage, which would also make them a quite important sector to consider (Petrecca 2007).

In the design process, consideration for the type of device (mobile or PDA), how it will be used (touch screen or keypad), and how the content is delivered (wireless, RFID or storage cards) will hugely affect consumer acceptance of the content and the technology involved. If any part of the exploration process or use of the device and /or content lacks an intuitive approach the consumer is likely to falter and have an incomplete or bad experience, resulting in a reluctance to use it again (Braiterman and Becker 2008).

The questions that have arisen when talking with managers / owners of heritage sites in previous interviews have been ‘will the visitor accept the use of these handhelds?’ and ‘will they know how to use them?’ (Wilson 2002, 2006). A few years ago, the answer would have been no, accept for a few who liked to be at the cutting edge of technology. This has changed quite considerably. In 2007, the smartphone industry increased by 44% to approximately 24.5 million units with the PDA units increasing by 54% in 2007 (Nielssen 2007) showing that acceptance is growing rapidly and becoming a recognised method for use within museum and heritage sites (Aoki and Woodruff 2000).This does not really seem to be where the issue lies for the overall success of these models, it lies more with the content and user behaviour.

In order to understand what content might be acceptable, it is necessary to research user behaviour, how they might like to use the content, what type of content is going to engage them and encourage them to explore further and how they will interact with this information on a small screen.

These questions form part of ongoing research for the current development of a prototype called The Talking Walls®. The Talking Walls® is a multimedia template application that is able to be adapted to suit any heritage site, museum or historical building. It has been designed for users to visually explore and navigate where they are visiting via a mobile phone/handheld device, or a kiosk/DVD and online. The Talking Walls® will enable visitors to learn about the history of a specific place / space / object and be able to see how this has changed over time. It is a form of digital story-telling (Miller 2004) with the history/context of the building or object forming the basis of the content, combining additional entertaining and educational elements to help provide a multimedia visually-rich resource.

A project funded by the UK Government through the South East Economic Development Association (SEEDA) is currently under development on the estate of Lord and Lady Montagu at Beaulieu, in the New Forest, Hampshire UK. The project involves the visual restoration of Beaulieu Abbey, a Cistercian Monastery that was partially destroyed during the Reformation, during the time of King Henry VIII in the 1500s. The exterior of the Abbey will be shown in 3D with the ability to 'fly around' the structures. The virtual tours of the monks' story, history, and visualisation of their lifestyles, will provide a guide for visitors to Beaulieu Abbey and be an educational tool for schools and colleges. The exterior of the Abbey will be shown in fully rendered three dimensional format with the ability to 'fly around' the structures using the mouse or pointing device of the computer equipment (hand-held, mobile, kiosk screen, desktop or laptop computer). It will be possible to enter the interior of the main buildings, 'walk through' them, look up, down and around, and zoom in to view specific features such as stone carvings, windows, doors, altar pieces and other fixtures.

It is also intended to develop an application suitable for PDAs and similar portable devices which will enable visitors to walk around the site and download the interactive information via a wireless network. For educational establishments, it will also be possible to access a virtual tour via a website. The software visualization of the Abbey is now complete and trials are planned to enable parties of schoolchildren and postgraduate students to tour the site using PDAs. The way in which they navigate the site and use the software functionality will be carefully documented. Discussions will be held to ascertain user reaction and attitude.

Educational Development for Children

The concept has been devised from the need to provide children with visually rich 3D material on which their imagination can feed and expand. As in the case of Dunster Castle, the visiting school children are faced with a fairy tale manor house set on a beautiful wooded tor a few miles from the sea. They are then asked to imagine this as a moor covered tor and hill fort with the sea lapping the tor base. This is quite a leap of imagination, but if given virtual images and animation of what it was like and to be able to 'walk-through' this environment, the task becomes easier and their imagination is fired up. This was the original task given by a volunteer guide in 1995, and forms the background to the development of the visually rich application to what is now called The Talking Walls®.

This need to provide children (and adults) images on which to feed their imagination is still relevant today. For thousands of years, storytelling has been the main method of learning, passing on knowledge from one generation to another either verbally, pictorially or written. Once the print process was invented, the same knowledge has been made into books, some beautifully illustrated to enhance and create emotion with the content. With the more recent personal computer technology, CD-ROMs and other forms of interactive content, games and puzzles became accepted methods of passing on this knowledge and engaging the user. Now with the Web v. 1.0 and 2.0. the information and knowledge base has exploded and is available to anyone who has a computer and / or mobile phone (Braiterman and Becker 2008).

For the majority of this time, people visited heritage sites initially to see how other people lived and learn a little of their culture (Poria, Butler et al. 2004) (Masberg and Silverman 1996). The information provided at these sites has increased considerably with the need to attract and educate potential visitors, creating learning environments as well as somewhere to relax and explore with family or companions on a day out. The consumer’s thirst for historical and cultural knowledge arrived hand in hand with the increase in technology and their subsequent expectation of being able to access information wherever they are, at any time (Hooft 2006).

There have been different methods given to feed this ‘thirst’ ranging from guidebooks (separate ones designed for children), to audio tours, static period displays, seek and find sheets, exhibitions, kiosk displays and more recently, web sites, some specifically aimed at children containing cartoon-like images, puzzles and games. These seem to be aimed primarily of a very young age range. The older child does not seem to be so well catered for, and although they might be interested in a more adult interpretation, the material on offer may not enthuse or involve them enough to learn more. An application that will allow any user to explore areas such as creative writing, art, mathematics, design and technology, geography and communities through which they will also learn about history and culture would be ideal. The Talking Walls® was designed with this as its prime goal and to also cater for mixed abilities and ages by being able to pick and mix from the content to suit their level and interest.

Involvement and Engagement

It has already been pointed out that a variety of users, from children to older adults, will be using The Talking Walls® applications in a variety of contexts and ways. The way they will want to use the technology will vary not just in terms of their age, but their interest in mobile devices and their desire to use them for learning purposes.

Firstly, the level and depth of involvement will vary. Laaksonen (1994) considers that there are three sides to involvement:

1. Cognitive perspective – is the perceived personal relevance of the item that is important to the consumer
2. Predisposition to act - the interest or level of emotional attachment, arousal, drive or motivation evoked by a stimulus – for example, the level of interest in the subject and the interest triggered by the potential for learning and variety that the user perceives in the application
3. Response view – the reaction of an individual to a stimulus, which affects the 'learn-feel-do' sequence

User interaction patterns will therefore vary according to a number of factors and it will be important for museums or galleries (hosts) to be aware of the wide range of use levels to which the application will be put, and that this will not just a simple case of age. Some users may need some initial hands-on training before they are familiar with the technology.

In an article for a seminar, Jennie Fordham, Education Officer at English Heritage/HHA wrote: "Learning is a process of active engagement with experience…Effective learning leads to change, development and the desire to learn more. It is all about involvement - if we are involved, we care, and this has to start with understanding… through learning at whatever age." (Fordham 2005)

Visiting an historic house or museum already places the user in a rich contextual environment. They can interact with their surroundings by looking, listening and touching. They can converse and learn from the objects and people around them but what will actually engage them and lead them to want to discover more either while they are still onsite or once they have left the site?

Interaction and entertainment are considered to be prime methods for engaging and involving users. Interactivity is what makes the digital media, whether on the PC, console or television very different from other media such as books, films and television programmes, it is the interaction of the user with the content that engages the user. The user is no longer being passive, listening to the audio or watching the video clip, they are actively changing their experience, making it their own, personalising and becoming involved. Each user creates their own experience, learning different skills and through choosing different characters, their storytelling imparts different information, or possibly the same as another, but in a way that is pertinent to their level of understanding and learning preference. This is the basis on which games are designed, although primarily as a form of entertainment, knowledge is still acquired through the story of the game.

Game playing has proven to be a great success with mobile phone users earning considerable revenue for the mobile industry. The ability to play a game at any time, any place and anywhere (snacking) helps in this success, but also the designers have discovered that designing games specifically for the mobile phone instead of 'shoe-horning' console games has increased in popularity becoming the 'cornerstone' of the mobile gaming industry (Loader 2006). This example perhaps gives an insight into how it might be possible to create a similar success for designing heritage content that can be used on a small screen. There may be a need for small 'snack' sized chunks (Moses 2007).

User Interaction

This creates considerable problems for the designer. No longer can the designer create content and dictate the path of the user, this would make it a passive experience; the user would have nothing to do but sit back and watch what happens, or at the very least, click the buttons indicated for another static image. There is not a great deal of difference between this and turning the page of a book or switching channels on a television; it is not interaction in its full sense. The success of Web 2.0 has proven that users want to generate their own content and are doing so extremely well, seemingly without help from designers. It is not certain whether, if the web site has not been designed in the right way, the user would be able to interact with it and post, create and re-generate their own content.

The designer has provided a template in which this can happen, as in YouTube's website. The question arises as to how designers can transpose this to a small screen and provide the user with a model they can interact with and re-generate content. A template needs to be created that will allow this to happen, providing access to a database of generic specific content which they can then re-model but through a clear, simple and easy to navigate 'small screen' design.

A choice of animated walk-through routes are also provided using the present or previous owners as guides to show you round their home, as in the Beaulieu Abbey project outlined above. These guides provide additional information about the rooms, spaces or objects the users are visiting and will make the virtual experience more personal by being tailored to their own interests. The walkthroughs will enable people with limited mobility to enjoy and participate in exploring a space, in areas they may not be able to easily visit. The user would be able to create their own points of interest as they navigate around the site, building a more personalised tour which they can expand upon at a later date, possibly online.

Dynamic Interface

The KubeMatrix™ is the interface to all of this content. It is a unique concept acting as a navigational tool, representing time and space, content and links or objects and connections in a 4-dimensional cube matrix structure, as the interface navigating through the content and allows the user to choose their age range and content. The user will then be able to track their progress, with the KubeMatrix™ indicating the "spaces" they have visited and the routes taken. Each cube of the KubeMatrix™ represents a room or object or other designated space, and the connecting tubes are the passageways, connections or pathways between. The vertical layers are the different eras that are important to the history of the site, or the different hierarchies or links to content or objects of that time. Users will be able to see how the different eras impacted on day to day life by jumping from a room in one era to the same room in another era.

Children will be able to grasp their historical background much more easily and have an element of fun in doing so. The cubes and layers can be used as content holders with the interconnecting links creating a visual map of how the content is related and how to navigate between it. The KubeMatrix™ will also work with GPS, allowing the client site to track the areas most (or least) visited, and record the user's visit. The recording could then be sold back to the user as a memento of their day, and can be stored for future visits. Should the user return, they would be able to track the areas already visited and explore new areas. They will be able to feed back to the database via the website.

The prototype has been tested with varied audiences ranging from individuals to small groups with good feedback and several comments on how easy it was to use. The templated interface has been designed to be simple, very visual and intuitive to use. This has been based on the web design general rule of thumb being three clicks to get to your destination within the site, this paradigm should follow for any screen based application, allowing easy and quick navigation through the content (Nielsen 2000).

Human Interface

Generally, theories around the diffusion of innovation seek to explain and predict adoption rates. Factors that have been found to influence adoption rates include:

Adopter characteristics

Social networks

Communication process

Promoter's strategies

Innovation attributes, e.g. usability

Constantiou, Damsgaard & Knutsen (2007) suggest that it is not only critical to identify core characteristics of adopter types, but also to determine the various categories of users based on their behaviour patterns and how they perceive a technology application might be of value. For example, a casual visitor to a museum may not wish to take the time to work out how a visual tour could be of maximum benefit, whereas someone with a keen interest in art might tour a gallery exhibition and take time to learn the application in order to extract the information and the benefits they require.

Constantiou, Damsgaard & Knutsen further believe that a learning perspective should be taken towards users and categorise mobile users according to four types: talkers (who use mobile voice applications only); writers; photographers and surfers. The adoption of technology and its services will evolve over time as people learn. Social networks are very important in this regard. Early adopters will discuss it with late and laggard adopters and educate each other about the possibilities afforded by the application.

In contrast, Sarker and Wells (2003) state that factors affecting the implementation and acceptance of wireless technology are interface characteristics and network capabilities. Both are crucial in any application in a heritage situation.

In particular, Sarker and Wells believe that key factors affecting usage are:

Age of the user-adopter

Technology self-efficacy, i.e. the individual's confidence in using devices

Cultural origin e.g. some users may view a visualisation as trivializing a serious subject

Interface characteristics – a device and its application must be easy to use

Network capability – reliability is essential, including immediacy and the ability to search quickly for the specific information required

Connectivity

With regards to connectivity, The Talking Walls® is designed to work via wireless technology or by downloading to a storage card prior to visiting the site. Both methods would allow the user to freely roam around the site and to have a degree of interaction with the content and the site or exhibits. This also allows the user to choose what they would like to explore and the type of content with which to do this. This is very important for the user, especially those that are used to mobile technology and like to control the amount of information they are interacting with. If they do not have a suitable phone or handheld device, the basic content can be supplied already downloaded to a wireless PDA or other handheld device which the user would hire on site. The additional content for this method would be able to be downloaded as they are exploring. They would then be able to purchase this storage card and continue to extend the experience, exploring in further detail chosen areas via the website on their home computer.

A popular method of providing information to the handheld device has been via RFID tags. An RFID tag is an object that can be attached to or incorporated into, a product which stores or remotely retrieves data regarding that object. RFID tags push the appropriate information to the user when they are in front of each item, such as a sculpture or painting. This links the user to being in front of the object until the information has been given. Not only this but they may also be impractical in some circumstances, i.e. a user may walk past an object and the information is fed to them even though they have not chosen to stop at this exhibit.

The user may also need to be reasonably near the exhibit for the information to flow, which again, may not be practical i.e. the exhibit is roped off or too high such as paintings on ceilings and borders (Aoki and Woodruff 2000). Most importantly, the information is a one way flow; interaction does not exist in this method, the user will not be able to generate their own information back to the tag, perhaps only back to the web site if the handheld device is wireless enabled.

From a user's perspective, this method may be the most frustrating with regards to content flow and interaction and not one to use as the main source of technology with the Talking Walls®. It may be that a combination of technologies will be the answer to providing a seamless interaction with the device, the site and the user. Any attempt at providing an application that is difficult to use, unreliable or does not provide the user-driven data promptly is unlikely to be popular. Users must feel confident that, as they find their way around a museum or gallery, the technology is a genuine learning aid that will greatly enhance the quality of their visitor experience, seamlessly.

Conclusions

The application of visualisation technology across a range of mobile and web-based platforms is beginning to be adopted by tourism, heritage or educational organisations where the need to provide educational content has to adapt to the opportunities afforded by modern technology and the ways in which young people are interacting with learning and information environments. the applications across a range of environments are still at the beginning of the technology life-cycle and involve not just development of hardware, software and web enabled domains but also an understanding of user interaction, engagement and involvement.

This paper has attempted to explore some of these issues in the context of actual development projects. Additional research will be required to ascertain the efficacy and effectiveness of these and future projects and trials of the Beaulieu Abbey site will offer the opportunity to assess the behavioral and attitudinal factors of visitors using PDA devices.

References

Aoki, P. M. and A. Woodruff (2000). *Improving Electronic Guidebook Interfaces Using a Task-Oriented Design Approach*. Museum and the Web 2000

Barry, A. (2005-6). *Engaging with a larger, more diverse audience.* Annual Review, Natural History Museum

Braiterman, J. and M. Becker (2008) "*Academic Review: Customer Experience and Mobile Marketing.*" Jared Research / iLoop Mobile

Constantiou, I., Damsgaard, J. & Knutsen, L. (2007). *The Four Incremental Steps Toward Advanced Mobile Service Adoption. Communications of the ACM, Vol. 50 No. 6*

Fordham, J. (2005). *What is learning all about - why does it matter*? HHA Seminar

Hooft, M. v. t. (2006). *Ubiquitous Computing: More than Handhelds.* HandHeld Learning 2006. London

Kardes,. F (1999). *Consumer Behaviour and Managerial Decision Making.* Addison Wesley Longman Inc.

Loader, J. (2006). *Entertaining Mobile. Mobile Entertainment Markets - Opportunities and Forecasts 2006-2011*, Juniper Research

Masberg, B. A. and L. H. Silverman (1996). "*Visitor Experiences at Heritage Sites: A Phenomenological Approach.*" Journal of Travel Research

Miller, C. H. (2004*). Digital Storytelling - a creator's guide to interactive entertainment.* Oxford, Focal Press, Elsevier.

Moll, c. (2005) "*Mobile Web Design: Methods to the Madness.*" Mobile Web Design,

Moses, A. (2007) "*'Snacking' to spark mobile video explosion.*" The Age

Nielsen, J. (2000). *Designing Web Usability.* Indianapolis, New Riders Publishing

Nielssen, B. (2007) *"Implementing a Mobile Strategy."* Business News Volume

Petrecca, L. (2007) *"Tech giants target older buyers — and their cash."* USA TODAY

Poria, Y., R. Butler, et al. (2004). "*Links between Tourists, Heritage, and Reasons for Visiting Heritage Sites*." Journal of Travel Research

PR Newswire (2007) "*M:Metrics Study: 92.5 Million Active SMS Users Make Short Code-Based Mobile Marketing the Most Effective Platform for Mobile Advertisers*." News and Information

Ranchhod, A. (2007). "*Developing Mobile Marketing Strategies*." International Journal of Mobile Marketing

Sarker, S. & Wells, J. (2003). *Understanding Mobile Handheld Device Use and Adoption*,Communications of the ACM, Vol 46 No. 12

Wilson, D. (2002-2006). *The Talking Walls research*. D. C.-N. T. Historic Houses Association, British Museum.

AN EXAMINATION OF MINISTERIAL LEVEL E-GOVERNMENT IN CARICOM STATES

Rhoda C. Joseph, Pennsylvania State University Harrisburg, (717) 948-6144, ruj1@psu.edu
Patrick I. Jeffers, University of the West Indies St Augustine, (515) 278-0414, Patrick.Jeffers@fac.gsb.tt

ABSTRACT

This paper focuses on the ministerial level of e-government in the Caribbean. The paper specifically focuses on the regional bloc of Caribbean countries called Caricom. Caricom nations have clearly defined goals in terms of using technology to increase the social, economic, and political well-being of its citizens. This paper uses a content analysis methodology to examine ministerial level e-government in the region. Government ministries are functional organizations responsible for assigned portfolios of the state government. The paper further addresses unique factors that affect the development and implementation of e-government projects for Caricom countries.

KEYWORDS: Caricom, Caribbean, e-government, ministerial e-government, content analysis

INTRODUCTION

E-government is a global phenomenon. The United Nations (UN) 2008 e-government report indicates that virtually all of the 193 states it recognizes have initiated some form of e-government project [11]. In its most elementary form, an e-government initiative involves a static website that allows users to access information about services provided by a government agency. In this paper, we focus on e-government development within the regional trading and collaborative bloc of Caribbean member states known as Caricom (Caribbean Common Market).

Our paper begins with a brief overview of the region and a discussion about the ministerial structure of government in Caricom states. Ministries are departments of the central government, typically classified into functional domains such as education, health, finance, agriculture, and tourism. While some e-government projects in developed nations have transformed into complex, integrative web 2.0 platforms, many developing regions are at early stages of development with static web-pages on their websites [6]. This paper highlights the concerns and challenges that persist in a global context as emerging and developing nations pursue the development of e-government infrastructure.

We use empirical data collected and analyzed by way of a website content analysis to examine the nature and level of e-government development among Caricom government ministries. Current e-government literature posits that citizens are the dominant focus for e-government projects [7] [12]. Consistent with this premise, the focus for e-government in the Caribbean is also the citizen, while other areas such as the business and employee foci are much less developed [5]. This paper specifically examines the citizen focus with an examination of services provided by e-government websites of several government ministries in the region. The primary research questions are: what is the level of development of ministerial e-government projects? For existing e-government projects, what e-services do government ministries provided and what is the impact of these e-government services on citizens?

CARICOM MINISTRIES AND E-GOVERNMENT

Fully leveraging the potential of an e-government project would require an understanding of the indigenous conditions under which it is constrained. The cultural, economic and political complexity of the Caribbean archipelago represents a particularly unique research opportunity for exploring the potential of e-government projects. British, Dutch, Spanish and French colonizers have left their political imprint, which, even today, continues to affect the political landscape of the region. Caricom was initially formed by a group of former British colonies after gaining their independence from colonial rule in the mid-20^{th} Century. Listed alphabetically the current members of Caricom are: Antigua/Barbuda, Bahamas, Barbados, Belize, Dominica, Grenada, Guyana, Haiti, Jamaica, Montserrat, St. Kitts/Nevis, St. Lucia, St. Vincent & the Grenadines, Suriname, and Trinidad and Tobago.

One of the key objectives of Caricom is the creation of a single market and economy, also referred to as CSME (Caricom Single Market Economy). The development and improvement of the information technology (IT) infrastructure, of which e-government is an integral part is linked closely to the “sustained economic development” goal to which Caricom aspires [3]. The governments of Caricom nations have readily embraced the new global developmental shift where knowledge management is critical to both economic and social development policies. Jamaica has developed an electronic transaction policy and enabled legislations with “the hope that an appropriate ICT policy will generate new products, new production processes and new forms of organization and competitiveness,” [8]. Trinidad and Tobago, having openly committed to transforming the country into First World status by 2020, has formally launched its National ICT Strategic Plan with similar aspirations for “empowering the people, innovation, education, information technology and infrastructure, thus creating an empowering environment for accelerated social, economic and cultural development” [13].

E-government provides an opportunity for ICTs to have a direct and potentially positive impact on the lives of users. In the Caribbean region, e-government websites serve an important role by informing citizens about current issues, including disease mitigation and prevention, educational opportunities, and disaster preparedness and recovery for natural disasters such as hurricanes and floods. The use of e-government tools and services by government ministries represents very practical applications that can ultimately improve the lives of Caricom citizens and the region as a whole.

As former British colonies, English speaking Caribbean countries have adopted modified forms of the Westminster style of government [4]. The Westminster parliamentary democracy is comprised of three branches of government: the Executive, the Legislature and the Judiciary. Of relevance to this discussion is the Executive branch of government. The Executive branch consists of the Prime Minister and Cabinet. Under the Prime Minister, there is a subset of ministers that are assigned to various ministries of government. As each nation develops its e-government footprint, resources may be disproportionately divided across ministries. This can be based on a variety of factors including overall country goals, Caricom initiatives, and localized champions of e-government projects.

RESEARCH FRAMEWORK AND METHODOLGOY

This section details the framework and methods used to answer the primary questions of this paper. A content analysis method was used for the data collection. Content analysis uses secondary sources of data such as articles, advertisement, websites and videos to collect data about a particular theme or subject. Recent studies have utilized content analysis for research in areas such as e-government development [5], intellectual capital disclosures [1], and case methods in supply chain management research [9]. The main strength of this methodology is that it is unobtrusive and relies on previously existing material. However, since the methodology does not utilize primary data sources such as interviews and surveys, it may not necessarily present a comprehensive picture of the subject matter being examined. Ultimately, no methodological approach is void of inherent strengths and weaknesses, and it is thus important to utilize methods that are most appropriate for the context of the study. E-government websites provide complex and relevant information that is appropriate for this content analysis study.

The first research question seeks to determine if ministerial e-government initiatives are present for Caricom nations. We first used the country’s main e-government portal to determine what ministries exist in that specific country. After finding the names of ministries, we then further searched for a unique website for each specific ministry. We then identified a set of ministries that were common across all countries and then proceeded to question two. The second research question examines what are the primary e-government services available at ministerial websites and impact of such services. We examined each ministerial website in detail, and compiled data about the sites main features and the services that it provided. We used the raw data to compile a summary of the primary services available at each of the different ministries for each Caricom country. Even though many studies have used stage models to classify e-government development [2] [6] [10], the extant literature is void of studies examining the development and impact of ministerial level e-government.

OUTCOMES AND ANALYSIS

The following five ministries were common across Caricom nations: agriculture, education, finance, health and tourism. In some instances there was missing data because we did not find an official government website for that particular ministry. We define an official site as a site that is connected to the country’s main e-government site, or explicitly states on the page that it is a branch of government or is a site that has the .gov extension in their web address. Due to the dynamic nature of the web, the countries with missing values may at the present time have a current site available. It is also possible that due to the methods that we used to search in this study, we did not find an existing site at the time that data was collected.

For the content analysis we visually searched each site to determine if specific items were available to users. For each item we tallied the number of sites found that contained that specific feature. From the content analysis two categories of features emerged. The first category was a set of static features that provided descriptive information about the government ministry. The primary descriptive features identified were: 1. Contact information, 2. Mission of Ministry, and 3. Hours of operation. The second category of features provided some degree of interactivity to the user. The main items identified under this second category were: 1. External links, 2, Email, 3. Search engines, 4. Downloadable forms, and 5.Message boards. The data was then combined to show the percentage of sites that contain a particular feature. For example, we found that 93% of ministry of agriculture sites had contact information available, while only 47% of ministry of health sites had contact information available (Table 1).

Table 1: Percentage of Ministerial Websites with Specific Features

Feature Available	Agriculture	Education	Finance	Health	Tourism
Static Features					
1. Contact Information	93%	85%	85%	47%	93%
2. Mission of Ministry	77%	54%	85%	54%	85%
3. Hours of Operation	24%	24%	8%	8%	31%
Interactive Features					
1. Email	85%	70%	77%	39%	93%
2. External Links	93%	93%	93%	62%	100%
3. Search Engine	31%	39%	39%	24%	54%
4. Downloadable Forms	24%	39%	24%	8%	39%
5. Message Board	0%	16%	0%	0%	31%

DISCUSSION

From the above results the two ministries with the most developed features available to users were agriculture and tourism. Agriculture and tourism represent a key form of revenue for many of the Caricom nations identified. It thus, seems more than co-incidental that these two ministries have received the most attention from e-government developers. The core of developing ministerial level e-government is to provide easy access directly to the individuals that need it. Both agriculture and tourism provide easy accessibility to users by having contact information and email features readily available. Every tourism website also provided users with external links. The external links included access to products and services provided by local businesses such as hotels, car rentals, tours, as well as international information such as travel agents and airlines.

Education and Finance are next with similarity in the content and quality of available features. These two ministries are core functional units in the government of Caricom countries. Education is stressed in the region, with most islands having literacy rates of over 90% for the adult population. Finally, of the ministries examined, the ministries of health seem to be the least developed in terms of the content and quality of on line services that are provided. This may be as a result of the more localized manner in which health care services are administered. Health care providers such as doctors and nurses operate in independent private practice environments or regionalized hospitals. Further investigation through interviews and field visits will be needed to acquire more useful information about the development of ministerial level e-government health initiatives.

Ministerial level e-government projects present opportunities for government to provide important services directly to constituents. For Caricom countries, citizens in one nation can easily access information about jobs, housing, and educational opportunities in another nation via an e-government platform such as a ministerial website. Effective e-government programs can facilitate easier access to employment, tourism, education, healthcare and even recreation for Caribbean citizens.

CONCLUSION

An important contribution of this paper is to expand the current, albeit limited literature available on e-government applications in the Caribbean region by specifically, looking at ministerial level e-government projects in the region. E-government in the Caribbean allows citizens to access ministerial information, allows Caricom nationals to reduce the physical divide, and may even bring ex-patriots closer to their home country. Ministerial level e-government, though narrower in focus that large scale country or federal level projects can have far reaching impact for users. Ultimately, the ideas presented in this paper, can provide a context for ministerial level e-government studies in other similar developing regions of the globe.

REFERENCES

[1] Abeysekera, Indra, "Intellectual capital disclosure trends: Singapore and Sri Lanka," *Journal of Intellectual Capital*, 2008, 9(4), 723-37.

[2] Andersen, Kim Viborg and Helle Zinner Henriksen, "E-government maturity models: Extension of the Layne and Lee model," *Government Information Quarterly*, 2006, 23(2), 236-48.

[3] Caricom Secretariat , "The CARICOM Single Market and Economy (CSME)," http://www.caricom.org/jsp/single_market/single_market_index.jsp?menu=csme&prnf=1, 2009, Date Accessed: March 13, 2009.

[4] Griffin, Clifford E., "Democracy in the Commonwealth Caribbean," *Journal of Democracy*, 1993, 4(2), 84-94.

[5] Joseph, Rhoda C. and Patrick I. Jeffers , "E-Government Maturity in the Caribbean Nations," *Journal of Global Information Technology Management*, 2009, 12 (1), 52-70.

[6] Layne, Karen and Jungwoo Lee, "Developing fully functional E-government: A four stage model," *Government Information Quarterly*, 2001, 18(2), 122-36.

[7] Marche, Sunny and James D. McNiven (2003), "E-government and E-governance: The Future Isn't What It Used to Be," *Canadian Journal of Administrative Sciences*, 2003, 20(1), 74-86.

[8] Paulwell, P. , "Presentation By Hon, Phillip Paulwell Minister of Commerce, Science and Technology" *International Institutie of Communications 35th Annual Conference*,

October 12, 2004. http://www.mct.gov.jm/IIC_speech.pdf Date Accessed: April 13, 2009.

[9] Seuring, Stefan A., "Assessing the rigor of case study research in supply chain management," *Supply Chain Management: An International Journal*, 2008, 13(2), 128-37.

[10] Siau, Keng and Yuan Long, "Synthesizing E-government Stage Models: A Meta-synthesis based on Meta-ethnography Approach," *Industrial Management & Data Systems*, 2005, 105(3/4), 443-58.

[11] UNPAN, "United Nations e-Government Survey 2008: From e-Government to Connected Governance," in Department of Economic and Social Affairs Division for Public Administration and Development Management. New York.

[12] Wei, Xiujian and Jun Zhao, "E-government: Citizens' requirement analysis in Chinese e-government," in *Proceedings of the 7th international conference on Electronic commerce ICEC '05*, 2005.

[13] Weston, B., " Address bv H.E. Mr. Bernard Weston**.** Permanent Representative of Trinidad and Tobago to the U.N. and Head of the Trinidad and Tobago Delegation to the World Summit on the Information Society First Phase" Geneva, December 10-12, 2003. http://www.itu.int /wsis/geneva/coverage/statements/trinidad/tt.pdf Date Accessed April 13, 2009.

A FRAMEWORK FOR RESEARCH IN INFORMATION SECURITY MANAGEMENT

Sindhuja Parakkattu, University of Toledo, (419)-530-5644, menon.sindhuja@gmail.com
Dr. Anand. S. Kunnathur, University of Toledo, (419)-376-5391, AKunnat@utnet.utoledo.edu

ABSTRACT

Information security is a critical issue concerning organizations round the globe. All organizations involve in information-handling activities and therefore it becomes increasingly important to organize, manage and disseminate information in a useful and secured manner. Extant research in information security has been mostly focused on technological controls to protect information from threats and vulnerabilities. The information security literature widely discusses the role of information systems (IS) and information technology (IT) in secured management of information. However, practitioners and academicians have started to realize that effective organizational information security lies in the coordination of people, processes and technology. This motivates the development of a research framework for information security management that ensures the selection of adequate and proportionate security controls that protect information assets and give confidence to business stakeholders. As organizations become more and more interconnected, an effective information security management will help to build trust and commitment in inter-organizational activities.

AN OVERVIEW OF INFORMATION SECURITY RESEARCH

In today's dynamic and competitive business environment, an effective information system is part of the essential infrastructure of most organizations. Information systems include not only the hardware, software, data and other information assets, but also the people, policies, and procedures associated with the gathering, distribution, usage and maintenance of the information. As organizations rely more and more on information systems to perform most of their business operations, concerns about controlling and securing information become paramount. Increased organizational dependence on information systems has led to a relative increase in the impact on the organization of compromised information security [1]. In this context, information security management (ISM) is a critical issue that is beginning to attract the attention of the communities of research and practice. ISM focuses on streamlining the management activities that creates an organizational framework within which the information system operates and mainly aims at protecting the information assets of the organization [2]. It includes ensuring the security of information through proactive management of information security risks, threats and vulnerabilities. This necessitates the need for ISM to be built into the daily business operations and alignment with the overall business objectives of the organization.

The real challenge of information systems is to ensure that the information is of highest quality in terms of timeliness, completeness, accuracy, confidentiality, reliability, readability and appropriateness [3, 4, 5]. As organizations experience unacceptably high levels of security abuses, they seldom provide consistently high quality information resources to meet manager's requirements [6]. The cost of compromising the information for any reason is extremely grave in terms of the damages caused due to monetary losses, disruption of internal processes and communication, loss of potential sales, loss of competitive advantage, wastage of time, efforts and manpower and even business opportunities, while it also damages the reputation, goodwill, trust and business relationships [7,8].

Most of the past studies on ISM focused on the technological [9] and administrative [10, 11] issues from an IS or IT perspective. However, the challenges faced by ISM stem from those related to the management of organization as a whole. In spite of the vast resources expended by organizational entities attempting to secure information systems through technical controls and restrictive formal procedures, occurrences of security breaches and the magnitude of consequential damage continue to rise. The weakest link in the security chain appears to be the absence or inadequate emphasis on the behavioral and organizational aspects of ISM. Effective organizational information security depends on managing the three components, namely; people, process and technology. Werlinger et al., [12] tried to provide an integrated view of human, organizational and technological factors that contributed to the complexity of security related challenges. The study aimed at providing suggestions for improving the security tools and processes. Though they have identified and described 18 challenges that can affect the ISM within an organization, the paper is silent on implications on organization's performance. Hagen et al., [13] tried to assess the effectiveness of implemented organizational information security measures and suggested that awareness creating activities should be encouraged in organizations where security measures are implemented. Though the authors looked at the effectiveness of such measures from a technical and administrative stand point, the study has not taken into consideration other critical factors of management. Further, implications of assessed effectiveness of security measures on organizational output are not dealt with. Studies have been done to measure the effectiveness of ISM from various individual dimensions. Chang and Lin [14] examined the influence of organizational culture on the effectiveness of ISM implementation. Authors suggested that human dimension of information security cannot be resolved by technical and management measures alone. They proposed a research framework relating organizational culture traits with the principles of ISM. Ashenden [15] addresses the human challenges of ISM and pointed out that information security management depends on technology, processes and people. Author suggests that organization should look into the skills that are needed to change the culture and build effective communication between all members of the organization, with regards to information security.

It is evident from the available information security literature that while ISM is a multidimensional phenomenon, reflecting technical, management and institutional perspectives [16], most of the research emphasis has been on the technical and formal aspects of ISM. Effective ISM seems to be an organizational challenge and no longer merely a technical commitment. In this regard, the research framework we propose to develop, examines the challenges of ISM by exploring the objectives, practices and other management factors that could influence the organizational performance and competitive advantage.

ISM Objectives and Practices

To safeguard organizational information assets from internal and external security threats, variety of information security standards and guidelines have been proposed and developed. The phrase "security framework" has been used in a variety of ways in the security literature over the years, but British standards (BS 7799) promoted the term information security management system (ISMS) and came to be used as an aggregate term for the various documents and architectures, from a variety of sources, that give recommendations on topics related to information systems security, particularly with regard to the planning, managing, or auditing of overall information security practices for a given institution. BS 7799/ISO 17799 deals with ISMS requirements and is used within companies to create security requirements and objectives. The Generally Accepted System Security Principles (GASSP) is a joint international attempt to develop a protocol to achieve information integrity, availability and confidentiality. However, ISO 17799:2005 (ISO 27001) is the widely accepted and suitable model for ISM, as it adequately addresses various security issues in organizations [17].

Qingxiong Ma et al. [18] examined the objectives of ISM and management practices used to achieve the same, as well as the relationship between information security objectives and practices. They identified four objectives which are most frequently considered for ISM. They are confidentiality, integrity, availability and accountability. Therefore, this proposed framework proposes to use these objectives for its purpose. ISO 17799 (ISO 27001) code of practice covers 10 control areas such as security policy, organizational security, asset classification and control, personnel security, physical and environmental security, communications and operations management, access control, systems development and maintenance, business continuity management and compliance. The authors refined these practices and obtained 8 commonly used practices by the ISM professionals. The framework also considers those 8 practices which is in alignment with ISO 17799 code of practice for ISM, as the basis for ISM practices.

Other critical organizational factors

Identification and addressing of other critical organizational factors that has practical significance to ISM will give a comprehensive perspective to the organizational view of information security management. As most of the operational, procedural and technical part of ISM is covered by the ISM objectives and practices, other factors that drive the need for ISM need to be considered. Based on the literature, some of the factors identified are top management support, organizational culture and structure, self-efficacy, and awareness creation [19].

Top Management Support: According to an Auburn University study, sponsored by the International Information Systems Security Certification Consortium ((ISC^2), obtaining senior management support is one of the most critical issues influencing information security effectiveness in organizations today [20]. The survey found that 62% of their daily tasks require the exchange of information or cooperation with others. And so implementing information security programs requires exceptionally high levels of task interdependence, which warrants greater levels of executive support to be successful. Knapp et al. [21] examined the impact of top management support on organization’s security culture and security policy enforcement. An organizational culture with less tolerance to good security practices is found with low levels of support and also retard the enforcement of security policies. Considering top management support to be an important driver for ISM, the study proposes to include top management support as one of its dimensions.

Organizational Culture: Culture is considered as the operating system of an organization, as it directs how employees think, act and feel [22]. It is also evident from the literature that culture paradigm is associated with the existing practices and roles in an organization [23]. Consequently, exploring the various cultural traits that facilitates an organization to perform ISM is of utmost importance from an organizational perspective. Hall [24] identified 10 streams of culture useful for addressing security issues that might emerge in any given setting. Later, Dhillon [25] named it as the web of culture consisting of 10 streams namely; interaction, association, subsistence, gender, temporality, territoriality, learning, play, defense and exploitation. Chang and Lin [14] used two dimensions, internal/external orientation and flexibility/control orientation, in their study on influence of organizational culture on ISM. The four constructs of organizational culture that emerged out these two dimensions were cooperativeness, innovativeness, consistency and effectiveness. The research framework proposes to use the Chang and Lin cultural constructs to measure organizational culture.

Self-efficacy: The eventual success of information security depends on appropriate information security practice behaviors by all who are associated with the system, and especially by the end users. Rhee et al. [26] explored the antecedents of individuals' self-efficacy beliefs in information security and tested relationships among self-efficacy in information security, security practice behavior and motivation to strengthen security efforts. This study also considers self-efficacy as an important construct for ISM in an organization.

Awareness Creation: Hagen et al. [13] pointed out that awareness creating activities have greater impact on ISM compared to technical and administrative measures applied by organizations. Increasing the awareness of security issues is the most cost-effective measure that any organization can envisage [25]. This framework considers awareness as part of the ISM dimensions.

Organizational Performance

Organizational Performance is a broad construct which captures what agencies do, produce, and accomplish for the various constituencies with which they interact. However, there is no universally recognized measure of organizational performance. Venkataraman [27] studied the perception of the respondents regarding organizational performance with respect to market and financial performance. This measure was used in many studies that examined the organizational performance [28, 29].

Competitive Advantage

When a firm's sustained profit pattern exceeds the industry average, the firm is said to possess a competitive advantage over its competitors. From a resource based perspective, a firm is said to have a competitive advantage when it is implementing a value creating strategy not implemented or not simultaneously being implemented by any current or potential player. It defines capabilities that differentiate an organization from its rivals. Suhong Li et al. [29], in their study used price, quality, delivery dependability, product innovation and time to market as the dimensions of competitive advantage construct.

Research Agenda

We represent the framework using the conceptual model given in figure. 1. The model depicts organizational factors to be the drivers of information security management. ISM objectives and practices are dimensions to assess ISM. Further, the influence of ISM, driven by the organizational factors, on the performance and competitive advantage is represented in the model. The research framework proposes to:

- Develop a comprehensive framework for ISM, reflecting, in addition, the organizational dimensions of security concerns.
- Examine the role of each dimension towards effective ISM.
- Examine the influence of ISM dimensions on Organizational performance and Competitive advantage

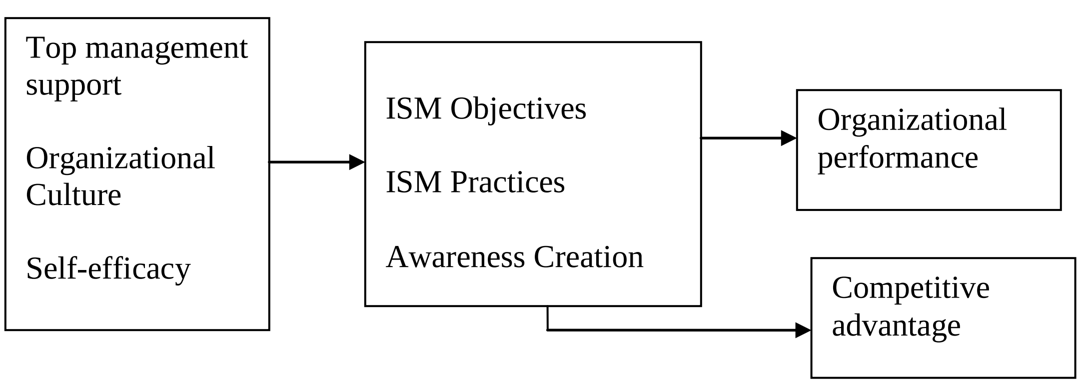


Figure 1: The Conceptual Model

Deliverables

Every business, big or small, faces major financial consequences due to loss of data or a breach of security. Out of the various types of security breaches happening in US, 47% accounted for the security incidents involving corporations and businesses [30]. At the bottom line, a business cannot afford to take the risk of ignoring data loss and security breach exposure. Therefore it is imperative that an organization give due consideration to the information security management aspects. This conceptual framework aims at providing a better understanding of the information security objectives and practices, considering other organizational factors, for an effective information security management. Information security management plays a vital role in addressing the security, compliance and efficiency needs of an organization. This provides a vast range of benefits which includes a holistic understanding of organizations' security status of the assets, prioritizing security occurrences, evading security breaches and demonstrating conformity with regulations in a much more efficient fashion than in the past.

We envision the developed framework to help:

- Explore approaches to integrate ISM within the organization
- Develop an information security strategy for the organization
- Create a pervasive information security culture
- Build trust and confidence in inter-organizational activities and processes to strengthen the supply chain.

References

1. Kankanhalli, A., Teo, H-H., Tan, B.C., Wei, K-K. "An integrative study of information systems security effectiveness,". *International Journal of Information Management*, 2003, 23(2), pp. 139-154.
2. Karyda, M., Kiountouzis, E., Kokolakis, S. "Information Systems security policies: a contextual perspective,". *Computers & Security*, 2005, 24, pp. 246-260.
3. Wang, R. Y., Strong, D.M. (1996), "Beyond accuracy: what data quality means to data consumers,". *Journal of Management Information Systems*, 1996, 24(4), pp. 5-34.
4. Caby, E. C., Pautke, R. W., Redman, T. C. "Strategies for improving data quality,". *Data Quality*, 1995, 1(1), pp. 4-12.
5. Miller, H. "The multiple dimensions of information quality,". *Information systems management*, 1996, 13(2), pp. 79-83.
6. Garg, A., Curtis, J., Halper, H. "Quantifying the financial impact of information security breaches,". *Information Management and Computer Security*, 2003, 11(2), pp. 74–83.
7. Dhillon, G., Moores, S. "Computer crimes: Theorizing about the enemy within,". *Computers & Security*, 2001, 20(8), pp. 715-723.
8. Bruce, L. "Information security – key issues and developments,". 2003, available at:www.pwcglobal.com/jm/images/pdf/Information%20Security%20Risk.pdf.
9. Siponen, M.T., Oinas-Kukkonen, H. "A review of information security issues and respective research contributions,". *The Database for Advances in Information Systems*, 2007, 38(1), pp. 60-81.
10. Kraemer, S., Carayon, P. "Computer and information security culture: findings from two studies", In the Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society. Human Factors and Ergonomics Society, Orlando, Florida, 2005, pp. 1483–1487.
11. Mouratidis, H., Jahankhani, H., Nkhoma, M. Z. "Management versus security specialists: an empirical study on security related perceptions,". *Information Management & Computer Security*, 2008, 16(2), pp. 187-205.

12. Werlinger, R., Hawkey, K., Beznosov, K. “An integrated view of human, organizational and technological challenges of IT security management,”. *Information management & Computer Security*, 2009, 17(1), pp. 4-19.
13. Hagen, J. M., Albrechtsen, E., Hovden, J. “Implementation and effectiveness of organizational information security measures,”. *Information Management & Computer Security*, 2008, 16(4), pp. 377-397.
14. Chang, S. E., Lin, C. “Exploring organizational culture for information security management,”. *Industrial management and Data Systems*, 2007, 107(3), pp. 438-458.
15. Ashenden, D. “Information Security Management: A human challenge?”. Information security technical report, 2008, 13, pp. 195-201.
16. von Solms, B. “Information security – the third wave?”. *Computers & Security*, 2000,19(7), pp. 615-20.
17. Dhillon, G., Backhose, J. “Current directions in IS security research: towards socio-organizational perspectives,”. *Information Systems Journal*, 2001, 11(2), pp. 127-53.
18. Qingxiong Ma, Johnston, A. C., Pearson, J. M. “Implementation security management objectives and practices: a parsimonious framework,”. *Information Management & Computer Security*, 2008, 16(3), pp. 251-270.
19. Siponen, M. “A conceptual foundation for organizational information security awareness”, *Information Management and Computer security*, 2000, 8(1), pp. 31-41.
20. "Managerial Dimensions in Information Security: A Theoretical Model of Organizational Effectiveness," available at http://www.isc2.org/auburnstudyAbout (ISC)2
21. Knapp, J. K., Marshall, E. T., Kelly Rainer, R., Nelson Ford, F. “Information security: management’s effect on` culture and policy,”. *Information Management & Computer Security*, 2006, 14(1), pp. 24-36.
22. Hagberg, R., Heifetz, J. Corporate Culture: Telling the CEO the Baby is Ugly,”. Hagberg Consulting Group, San Mateo, CA, 1997, available at: www.hcgnet.com/research.asp.
23. Allen, D.K., Fifield, N. “Re-engineering change in higher education”, *Information Research*, 1999, 4(3).
24. Hall, E. T., *The Silent Language*, 2nd ed. New York, Anchor Books, 1959.
25. Dhillon, G., *Principles of Information systems Security*, NJ, John Wiley & Sons, 2007.
26. Rhee, H., Kim, C., Ryu, Y. U. “Self-efficacy in information security: Its influence on end users' information security practice behavior,”. *Computers & Security*, 2009, 28.
27. Venkatraman, N. “Strategic orientation of business enterprises: the construct dimensionality and measurement,”. *Management Science*, 1989, 35(8), pp. 942-962.
28. Croteau, A., Bergeron, F. “An information technology trilogy: business strategy, technological deployment and organizational performance,”. *Journal of strategic information systems*, 2001, 10, pp. 77-99.
29. Suhong Li, Ragu-nathan, B., Ragunathan, T. S., Rao, S.S. “The impact of supply chain management practices on competitive advantage and organizational performance”, *Omega*, 2006, 34, pp. 107-124.
30. Bennet, K. “The real risks of business”, retrieved from http://www.connecticutbusinesslitigation.com/tags/security-breach/.

Exploring E-Commerce assimilation gap in small firms: A knowledge management capability perspective

Neset Hikmet
Nicholls State University

Huseyin Zeybek
Univ. of South Florida, St. Pete

Maling Ebrahimpour
Univ. of South Florida, St. Pete

The independent variables are:

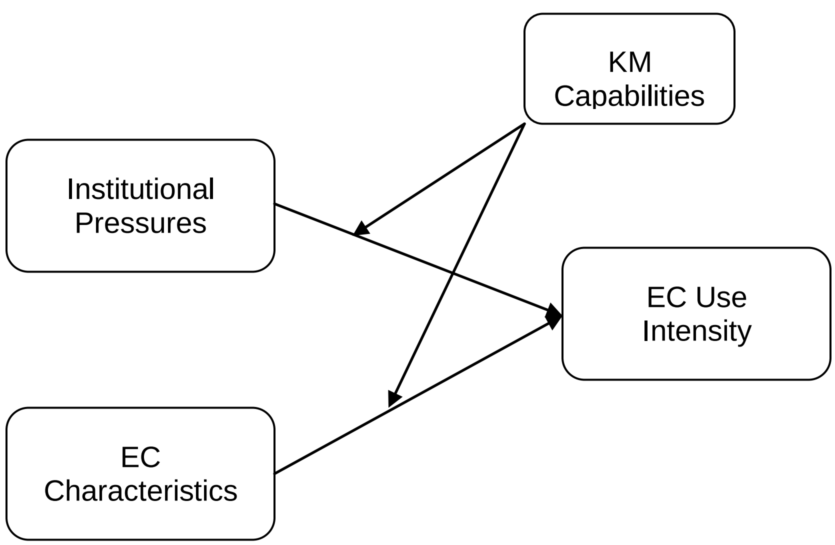
1) Institutional pressures: Mimic, coercive and normative pressures
2) EC characteristics: Relative advantage, Compatibility, complexity

The dependent variable:

Percentage and volume of EC use (similar to IJEC Hsu et al., 2006 and others)

Moderating variable KM capabilities:

Routines for acquire, analyze, integrate, and apply knowledge



Paths:

IP	➔	ECui	0.351	$p < 0.001$
ECc	➔	ECui	-0.166	$p < 0.01$

Rsq = 0.323

Moderating:

KMC	➔	IP	-0.113	NS
KMC	➔	ECc	-0.243	$p < 0.01$

The Influence of Foreign Demand on US Interest Rates

James Winder, Rutgers Business School, (732) 445-2996, jpwinder@rci.rutgers.edu

Abstract

The share of Treasury debt held by foreign investors has increased significantly over the last 20 years. We estimated an equation attempting to identify the effects of foreign demand for Treasuries on US interest rates. And we estimated another equation attempting to identify which central banks have had the largest effect on US interest rates. Our results suggest that when foreign central banks increase their holdings of Treasury securities, US interest rates decline. And, surprisingly, our results suggest that G-7 demand for Treasuries has had a larger effect on US interest rates than the purchases of the BRIC countries.

Introduction

The share of US Treasury debt held by foreign investors, both central banks and private investors, has increased substantially during the past 15 years. The share of Treasury debt held by all foreign investors moved in a range of 18% to 20% during the 1989-1994 period. The share increased to 52% during the fourth quarter of 2008. The latest data show it was about 49% at the end of the third quarter of 2009. Foreign central banks were primarily responsible for this increase. The share of Treasury debt held by foreign central banks increased from 12% in 1994 to 37% in 2008. Private foreign holdings of Treasuries, as a share of debt outstanding, have been steady since 2004.

There have been two general opinions on these trends. Financial markets analysts have for years focused on the vulnerability of US financial markets to any slowing in the pace of foreign demand. Their fear is that real interest rates would increase, boosting nominal US interest rates even if the rate of inflation is low. US Federal budget deficits and the increasing use of the euro as a reserve currency could cause foreign investors to become less willing to lend the Treasury funds. In a balance of payments framework, the US current account deficit has been easily funded by a strong demand for dollar-denominated assets, in their view. But, their concern is that the price at which the current account deficit is financed can increase. The result would be a slower trend in US economic growth.

The alternative view is shared by Federal Reserve chairmen Alan Greenspan and Ben Bernanke. Neither argues that the dependence on foreign investment is not a threat to long-term economic growth in the US. But they attribute low US long-term interest rates to a “global savings glut,” which they argue helped reduce real long-term interest rates in the middle years of the 2000-2009 decade. This savings glut, in their opinion, helped finance the accelerating US private sector demand for credit, which ended with the mortgage crisis. In a balance of payments framework, the strong ex-ante foreign demand for dollar-denominated assets has contributed to the current account deficit.

Plan of Study

This study will try to quantify the effects of increased foreign investment on the structure of US interest rates, and it will attempt to answer some questions that are relevant to the Fed's response to the financial crisis. Our estimation will cover the period from 1990 to the present.

Some of these questions include:

1.) How would a marginal shift in central bank demand affect US interest rates?

2.) So much of the demand for Treasuries over the past two years was a flight to quality. What might happen to US interest rates once the global credit crisis ebbs? The Fed will unwind its special portfolio activities, and foreign demand for Treasuries will diminish. What can we say about US interest rates when these two events unfold?

Hypothesized Equations

Our first equation attempts to explain Treasury yields as a function of expected inflation, Treasury supply, and security demand. This is essentially a Fisher equation where security supply and demand determine the real rate of interest.

Expected inflation will be approximated simply as a distributed lag on actual inflation. Demand variables will be represented by the ratio of a sector's holdings of Treasuries to the outstanding marketable stock. The equation will be estimated in log form, so the coefficients can be interpreted as elasticities. Our source for figures on holdings of Treasury securities will be the Federal Reserve's Flow of Funds data.

A general form of this equation might look like:

$$\text{Ln(Yield)} = c + \beta_0 \ln(\text{CPI}) + \beta_2 \ln(\text{CommB/Treas}) + \beta_3 \ln(\text{Fed/Treas}) + \beta_4 \ln(\text{ROW/Treas}) + \beta_5 \ln(\text{HH/Treas}) + \beta_{6\ldots n} \ln(\text{Other Sectors/Treas}).$$

Yield = Treasury constant maturity series for various maturities.
CPI = A distributed lag on the US Consumer Price Index, all urban consumers.
CommB = Commercial bank holdings of marketable Treasury securities.
Treas = The stock of marketable US Treasury debt outstanding.
Fed = Federal Reserve holdings of marketable Treasury debt.
ROW = "Rest of the World" sector holdings of marketable Treasury debt.
HH = Household holdings of marketable Treasury securities.
Other Sectors = All other sectors that hold marketable Treasury debt.

Our analysis will assess the influence on US Treasury yields of all sectors that hold Treasury obligations. If the coefficient on a ratio variable, for example ln(ROW/Treas), is statistically significant and negative, then we can conclude that an increase in the share of Treasury debt held by that sector contributes to a decline in nominal and real Treasury yields.

Assuming, for example, that the coefficient on the ln(ROW/Treas) ratio is negative and significant, then the issue becomes what might be the likely path of foreign demand over the next few years, when the federal deficit is expected to be over $1 trillion per year. The latest Flow of Funds data indicate that foreign holdings of US Treasury securities increased only slightly in the second quarter of 2009, after increasing very sharply in the first quarter. The possibility that foreign demand for Treasuries might be slowing makes our analysis very timely.

Our second equation estimates a foreign central bank demand function for Treasuries. It is possible to use IMF data on central bank holdings of foreign currency reserves to construct regional series of central bank reserve holdings. We will construct these regional series for the Euro Zone, Latin America, BRIC countries, the G-7, the Middle East, and Africa. These reserve pools can be used to explain central bank demand.

A general form of this equation might look like:

$$\text{Ln(FCB)} = c + \beta_0 \ln(\text{Euro Zone}) + \beta_2 \ln(\text{Latin America}) + \beta_3 \ln(\text{BRIC}) + \beta_4 \ln(\text{G-6}) + \beta_5 \ln(\text{Asia}) + \beta_6 \ln(\text{Middle East}) + \beta_8 \ln(\text{Africa}).$$

FCB = Foreign central bank holdings of Treasuries, using Flow of Funds data.
Euro Zone = Euro Zone central bank foreign currency reserve holdings.
Latin America = Central bank reserve holdings for Brazil, Mexico, Argentina, and Chile.
BRIC = Central bank reserve holdings for Brazil, Russia, India, and China.
G-7 = Central bank reserve holdings for the US, Canada, the UK, France, Germany, Italy, and Japan.
Asia = Central bank reserve holdings for Korea, Taiwan, Hong Kong, Philippines, Indonesia, Singapore, Malaysia, and Thailand.
Middle East = Central bank reserve holdings for the Middle East region, as defined by the IMF in the International Financial Statistics.
Africa = Central bank reserve holdings for Africa, as defined by the IMF in the International Financial Statistics.

This approach to estimation will permit some inferences about portfolio diversification following the introduction of the euro on January 1, 1999. We can test for demand elasticities before and after the introduction of the euro. What is exciting about a method that allows us to analyze the effects of the introduction of the euro is that we get away from the recycling argument. That is, the belief that the US current account deficit creates a natural demand for Treasuries that is automatically recycled back into the US asset market regardless of the spread between the returns on US and non-US assets. The proposed approach using IMF data puts portfolio choice right into the analysis, so there can be an effect on US interest rates. In order to provide more degrees of freedom, the estimation period will be extended back to 1990. We will estimate demand elasticities before and after the introduction of the euro.

Preliminary Results

The Equation for US Interest Rates

Our expectation is that many sectors will have a positive coefficient in the US interest rate equation. A positive coefficient indicates that as a sector's holdings of Treasuries increases, relative to the supply of marketable debt, US interest rates increase. A positive sign reflects those sectors, especially households, that respond to changes in interest rates. Their behavior does not represent a major influence on the level of interest rates. The behavior of these sectors does not determine the level of interest rates. These sectors would be excluded from a final estimation because of their limited effect on the level of rates. Rather, we are searching for those sectors that have negative coefficients. That is, when their share of Treasury debt increases, interest rates decline.

The general model discussed above did not yield successful results. Estimating the equation using many sector shares of Treasury debt outstanding as regressors produced few coefficients that were statistically significant. There was significant serial correlation in the residuals, which was reflected by Durbin Watson statistics well below 1.0. But, multicollinearity was not a significant problem between many of the sector holdings of Treasuries when measured as shares of marketable debt outstanding.

Because of the problems just discussed, the model was estimated using just one sector share at a time as a regressor to identify those sectors that can change the level of interest rates (negative coefficients) and those sectors that respond to changes in interest rates (positive coefficients). A distributed lag on CPI inflation is used as a proxy for inflation expectations. The 3-month Treasury bill yield is also included to capture changes in monetary policy and shifts in the yield curve. We also tested using sector holdings and the supply of marketable Treasury debt outstanding as separate regressors, instead of using the ratio term. Sectors with positive coefficients include commercial banks and households. Sectors with negative coefficients include the foreign sector and the Federal Reserve.

The estimate for the US interest rate equation is presented in Table 1. The yield on Treasury bills, expected inflation, and the supply of Treasury debt all have positive signs, while foreign central bank holdings of Treasury debt has a negative sign. As foreign central banks increase their holdings of Treasuries, market interest rates decline.

Foreign Central Bank Demand for Treasuries

The regressors in the foreign central bank demand equation show significant multicollinearity. The correlations between the currency reserve holdings of the BRIC countries, the currency reserve holdings of the G-7 countries, and the supply of marketable Treasury debt are shown in Table 2. The correlations are calculated for data in log form and in the first difference of logs. The differencing removes the multicollinearity, and so the model was estimated using the differenced data.

The reserve holdings of the BRIC countries and the G-7 nations have significant positive coefficients. Surprisingly, the results suggest that G-7 purchases of Treasuries have had a larger effect on US interest rates than have the purchases of the BRIC countries. The large and significant positive coefficient on the supply of Treasury securities suggests that there is a strong ex-ante demand for US debt. See Table 3 for the estimation results.

The Chow Breakpoint test was used to identify a structural change in foreign central bank holdings of Treasury securities. Various breakpoints were tested, and, surprisingly, the test revealed a break around 1990, and no further structural shifts after that. According to the Chow test, the introduction of the euro on January 1, 1999 did not cause a statistically identifiable break in the series. The Chow test results are presented in Table 4.

Conclusions

Our goal is to assess the effects of foreign central bank buying of Treasuries on US interest rates, and to analyze how the demands of regional central bank groups influence the level of interest rates. Our results suggest that foreign central bank demand has contributed to a reduction in US interest rates, and that the investments of the G-7 countries have had a larger effect on US interest rates than has the buying of the BRIC country central banks. Test results suggest that the introduction of the euro did not cause a significant change in central bank buying of Treasuries.

Summary of Results

Table 1.
Dependent Variable: LNT3Y
Sample Period: 1990:1 to 2009:1

Variable	Constant	LNTB3M	D(CPIEXP)	LNTSYOFF	LNTREAS
Coefficient	-462.4	0.40	29.18	-36.48	68.75
t-Statistic	-1.80	16.33	2.70	-4.27	2.99

Adjusted R-squared:	0.885	F-statistic:	147.37
Durbin Watson	1.30	Prob(F-statistic)	0.000
Akaike Info Criterion	8.32		

LNT3Y = the natural log of the yield on 3-year Treasuries.
D(CPIEXP) = the change in the natural log of a distributed lag on the CPI.
LNTSYOFF = the log of foreign central bank holdings of US Treasury securities.
LNTREAS = the log of the stock of marketable Treasury debt outstanding.

Table 2.
Correlation Matrices

	LNBRIC	LNG7	LNTREAS
LNBRIC	1.00	0.97	0.90
LNG7	0.97	1.00	0.94
LNTREAS	0.90	0.94	1.00

	D(LNBRIC)	D(LNG7)	D(LNTREAS)
D(LNBRIC)	1.00	0.01	-0.03
D(LNG7)	0.01	1.00	-0.02
D(LNTREAS)	-0.03	-0.02	1.00

LNBRIC = the log of the foreign currency reserves of Brazil, Russia, India, and China.
LNG7 = the log of the foreign currency reserves of the G-7 countries.
LNTREAS = the log of the stock of marketable Treasury debt outstanding.
D = the first difference.

Table 3.
Dependent Variable: D(LNTSYOFF)
Sample Period: 1990:1 to 2009:1

Variable	Constant	D(LNBRIC)	D(LNG7)	D(LNTREAS)
Coefficient	0.01	0.10	0.21	0.84
t-Statistic	1.47	2.19	2.54	4.70

Adjusted R-squared:	0.29	F-statistic:	10.66
Durbin Watson	2.09	Prob(F-statistic)	0.000
Akaike Info Criterion	-4.18		

LNTSYOFF = the log of foreign central bank holdings of marketable US Treasury debt.
D = the first difference.

Table 4.

Dependent Variable: D(LNTSYOFF)
Sample Period: 1980:1 to 2009:1
Chow Breakpoint Test: 1990:1

F-statistic:	68.68
Prob(F-statistic):	0.000

References

[1] Flow of Funds data, Board of Governors of the Federal Reserve System.

[2] International Financial Statistics, International Monetary Fund.

OPTIMAL PORTFOLIO ADJUSTMENT WITH CHANGING EXPECTED RETURNS

Peter M. Ellis, Department of Management, Utah State University, 84322-3500
Peter.ellis@usu.edu, 435-79702372

ABSTRACT

Linear programming models and their extensions have enjoyed a long history with portfolio optimization problems. Maximization of expected returns is typically sought under a set of restrictions. After an optimal portfolio is identified, sensitivity analysis of the objective function coefficients (the expected returns) might be of interest. The problem with this is that this sensitivity analysis traditionally must be done for a single coefficient at a time. There has not heretofore been known a method that permits this sensitivity analysis when several coefficients vary simultaneously. A new procedure is presented here that carries out the range of optimality sensitivity analysis for the optimal portfolio when expected returns are subject to synchronized and simultaneous fluctuations.

INTRODUCTION

There is a lengthy history of the application of linear programming and its extensions to the portfolio selection problem in finance. Objective functions might be either maximization or minimization. The goal might be to maximize expected return on the portfolio. It might be to minimize some measure of portfolio risk. Variations include maximin and minimax satisfying. Goal programming has been used where multiple objectives are desired. Markowitz [1952] very early formulated the problem as the minimization of a quadratic risk function with side restrictions. We now describe it as a quadratic programming problem and solve it with special adaptations of linear programming.

Linear programming has been a commonly used procedure in portfolio selection research and practice for many years. Speranza [1993] presented a variety of linear programming formulations related to portfolio selection. Gennotte and Jung [1994] introduced the additional consideration of including transaction costs. Kellerer et al [2000] extended that even farther by also including minimum transaction lots. Li and Deng [2000] also developed a linear programming portfolio model that included transaction costs. Konno and Yamazaki [1991] replaced the Markowitz quadratic risk approach with a model of absolute deviation of returns from the expected value, then used linear programming to obtain an optimal solution. This was followed by Chiodi, et al [2003], who also used absolute deviation to include portfolio risk. Ogryczak [2000] used goal programming for the portfolio selection problem with multiple criteria. Young [1998] sought a minimax portfolio selection solution with linear programming.

These linear programming approaches to the portfolio selection problem have examined problem formulation while including such ideas as risk, transaction costs and multiple investment goals.

There has never been great attention given to post-optimality analysis, especially objective function sensitivity analysis. The probable reason for that is that expected returns found in the objective function are quite likely to vary in some simultaneous fashion in accordance with market conditions.

In linear programming the optimal solution is open to further investigation through range of optimality sensitivity analysis. In brief, it is known that if an objective function coefficient is changed too much either positively or negatively, the optimal solution may change to a different extreme point. In terms of the investment problem, this means that a different portfolio mix might be called for. The analysis has always had the serious limitation that just one objective function coefficient change at a time can be considered. It has always been required that all other parameters of the problem be fixed and set while just the single objective function coefficient is permitted to vary. Of course, this is quite unrealistic in the portfolio selection problem with an objective of maximization of expected returns because the various individual returns on the candidate securities are very likely to fluctuate quite synchronously and in a predictable fashion. Changing market conditions cause security prices to change in response. The degree of commonality of response is heightened if the securities under consideration are closely tied to general market conditions.

Traditional linear programming sensitivity analysis permits just one coefficient change at a time when discovering the point at which the optimal solution will shift. Schenkerman [1997] extended the existing analysis by showing how to carry out a sensitivity study when several objective function coefficients vary simultaneously. That work solves the important case where the simultaneous changes are in constant ratio to each other. The development has remained in the theoretical world of mathematical programming, and there are no known examples of its application. However, it has great potential importance in situations where objective function coefficients actually do vary together. In fact, that is precisely the situation with the portfolio selection model.

SENSITIVITY ANALYSIS OF THE EXPECTED RETURNS IN THE OPTIMAL PORTFOLIO

We consider a typical linear programming portfolio optimization problem where the objective is to maximize expected return on the overall investment. The constraints serve to force diversification preferences. For the purpose of the work developed here, the model is quite basic, not including such advanced considerations as short selling or transaction costs.

The linear programming formulation is:

$$\text{Minimize } Z = \sum_{i=1}^{n} C_i x_i \tag{Ia}$$

Subject to:

$$\sum_{i=1}^{n} w_i x_i = 1 \tag{Ib}$$

$$\sum_{i=1}^{n} a_{ij}\, x_i\ (\leq)(=)(\geq)\ R_j \qquad j = 1,2, \ldots , M \tag{Ic}$$

$$x_i \geq 0 \tag{Id}$$

n = number of securities being considered for the portfolio
j = constraint number for any of the diversification requirements
C_i = expected return on security i
w_i = dollar percentage of the portfolio that is invested in security i
a_{ij} = risk diversification coefficient for security i in risk requirement j
R_j = limit on risk specification j
x_i = percentage of the dollar amount of the portfolio to be held in security i
M = number of constraint conditions upon the portfolio

The objective function (Ia) maximizes total percentage return of the portfolio. Constraint (Ib) requires that the dollar amount percentages of each security in the portfolio sum to 100%. The several constraints of the form (Ic) require that the risk and diversification requirements be satisfied. Last, the percentage allocations are required in (Id) to all be nonnegative.

The optimal solution identifies the dollar percentage allocation of each security to be included in the portfolio. After the solution is obtained, attention turns to the question of how the portfolio might change in response to different expected returns on any of the candidate securities. That inquiry has always been addressed by the postoptimality analysis known as the range of optimality. Briefly said, this sensitivity analysis shows an interval for each expected return coefficient C_i wherein the optimal percentage portfolio allocations stay the same. The analysis requires that all other parameters of the problem remain fixed, including all other C_i values. Of course, the total expected return would change in consequence of the new C_i value for security i. The criticism of this has always been that the expected returns C_i are very unlikely to singly and independently vary over time. It would typically be found that expected returns are all subject to the same market conditions and influences, and would thereby vary in relative synchrony.

A method for carrying out the range of optimality analysis when the expected returns vary simultaneously will be shown here. It uses the conventional sensitivity report found in all commercial linear programming packages. The procedure is limited to the important special case where the changes in the individual returns vary in constant ratio. One of the securities is chosen as the basis of comparison for fluctuations in expected return of the other securities. A beta coefficient b_i is calculated for every other expected return in comparison the expected return of the basis security. The b_i value = 1 for the basis security. Let D be the total expected change in expected return on the portfolio that arises from a 1% change in the expected return from the basis security. We have

$$D = \sum_{i=1}^{n} b_i x_i$$

A new constraint is formed by writing this as

$$D - \sum_{i=1}^{n} b_i x_i = 0 \qquad \text{(Ie)}$$

D is implicitly included in the objective function because it has a zero coefficient there. Also, D is not bound in any way in the constraints, so its presence in the problem does not alter the outcome. However, it is important to include for the purpose of carrying out the range of optimality analysis. Because D is a variable of the problem, its range of optimality appears in the conventional linear programming sensitivity analysis report. An upper bound and a lower bound on the objective function coefficient of D will be reported in any commercial linear programming software package. The bounds represent the maximum amount by which the expected return of the chosen basis variable can be permitted to increase or decrease without forcing the abandonment of the current optimal allocation and the consequential emergence of a new optimal allocation. Since the several expected return coefficients are linked through the associated beta coefficients, the result is that the sensitivity report for D shows the largest magnitude that each of the expected return coefficients can increase or decrease and still leave the current optimal portfolio unchanged.

AN EXAMPLE

Five ishares are being considered for the portfolio: S&P500 Index, S&P Midcap Index, S&P Smallcap Index, S&P Value Index and S&P500 Growth Index. Monthly return data are from the 60 months from January 2004 to December 2008. The basis security is selected to be the S&P 500 Index. The other four securities are made the dependent variable as a linear regression is run for each with the S&P 500 Index serving as the independent variable in each case. The beta coefficient is 1.00 for the S&P 500 Index, 1.2146 for the Midcap Index, 1.1955 for the Smallcap Index, 1.0165 for the Value Index and .9850 for the Growth Index. It does seem anomalous that the growth index was less than one over this time period. The respective R^2 values were 1.00, .9013, .8172, .9455 and .9468. It thereby seems generally true that expected monthly returns in all the securities vary synchronously and in direct ratio to changes in the monthly return on the S&P 500. The last returns available from these securities were for December 2008. They were .009, .03878, .0488, .008983 and .009015, respectively. These returns will be used as the C_i values.

We seek to maximize the total dollar value of expected return on a portfolio that consists of these five securities. No security can have a dollar value in excess of 30% of the portfolio. The sum of the S&P 500 and the S&P Value Index portions must be at least as large as the sum of the S&P Smallcap Index and the S&P Growth Index portions. The weighted average beta of the portfolio must be between 1.10 and 1.15. The linear programming formulation of the problem is shown below. The LINDO linear programming software of Schrage [1991] was used.

```
 MAX     0.009 SP + 0.03878 MID + 0.0488 SMALL + 0.008983 VALUE
      + 0.009015 GROW
 SUBJECT TO
   WEIGHT)   SP + MID + SMALL + VALUE + GROW =    1
     MIX1)   SP <=   0.3
```

```
    MIX2)   MID <=   0.3
    MIX3)   SMALL <=   0.3
    MIX4)   VALUE <=   0.3
    MIX5)   GROW <=   0.3
 DIVERSE)   SP - SMALL + VALUE - GROW >=   0
   BETA1)   SP + 1.2146 MID + 1.1955 SMALL + 1.0165 VALUE + 0.985 GROW
     <=   1.15
   BETA2)   SP + 1.2146 MID + 1.1955 SMALL + 1.0165 VALUE + 0.985 GROW
     >=   1.1
    SENS) - SP - 1.2146 MID - 1.1955 SMALL - 1.0165 VALUE - 0.985 GROW + D
     =   0
END
```

The variables are, respectively, SP, MID, SMALL, VALUE and GROW. The several constraints are named in accordance with their purpose. The constraint SENS) has the form of (Ie). It establishes D as the total change in the expected return on the portfolio when the return on SP changes by ±1% .

The optimal solution is:

OBJECTIVE FUNCTION VALUE

1) .29873900E-01

VARIABLE	VALUE	REDUCED COST
SP	.300000	.000000
MID	.300000	.000000
SMALL	.300000	.000000
VALUE	.050000	.000000
GROW	.050000	.000000
D	1.123105	.000000

Also, the range of optimality sensitivity analysis is:

RANGES IN WHICH THE BASIS IS UNCHANGED:

	OBJ COEFFICIENT RANGES		
VARIABLE	CURRENT COEF	ALLOWABLE INCREASE	ALLOWABLE DECREASE
SP	.009000	INFINITY	.000017
MID	.038780	INFINITY	.029781
SMALL	.048800	INFINITY	.039785
VALUE	.008983	.000017	INFINITY
GROW	.009015	.039785	.000032
D	.000000	.001016	.139261

Table 1 shows the expected returns on the several securities, as well as the resulting overall portfolio expected return. We denote this original problem formulation and solution as stage 1. Table 2 contains the optimal portfolio for each stage. For stage 1 the optimal portfolio weights are, respectively, .3, .3, .3, .05 and .05 The expected overall return on the portfolio is 2.987%. D has an objective function coefficient of zero. The sensitivity report in table 3 shows that the maximum increase is .001 and the maximum decrease is -.139. That is, the existing objective function coefficient of D is zero. The portfolio will not change unless that coefficient drops below -.138 or rises above .001. We first consider the case of a decrease.

Let the change in the expected return of the basic variable SP to be d. For each other security i the beta coefficient is b_i. The revised expected return coefficient for security i is then

$$C_{i\ new} = C_i + d*b_i \qquad \text{(II)}$$

Let the case of a negative d value be denoted stage 2. When d < -.139, (2) shows that each expected return coefficient would be negative. Because this is unattractive, the process is terminated. The other possibility is that there is an increase in the objective function coefficient of D. We denote this case as stage 3.

Because the range of optimality for the D coefficient has a maximum increase of .001, we set d to .002. This will be sufficient to bring about a new portfolio. Apply the beta coefficients to (II) and calculate new expected returns on each of the securities. For SP the new coefficient is .009 + 1*.002 = .011. For MID the new expected return is .03878 + (.002)*(1.2146) = .0412. The others are done in the same fashion. Table 2 shows the new values of the expected returns for each security. A linear programming model is created using the new expected return coefficients. The optimal portfolio for stage 3 is seen in table 1 to be different from that of stage 1. The expected return on the portfolio has risen to 3.212%. Note from table 3, the sensitivity report for D, that the range of optimality is 0 to infinity. Therefore, any increase in the magnitude expected return of the basis security SP will leave the portfolio mix unchanged from the distribution of stage 3. Naturally, the overall portfolio return will be increased, due to the increased expected return on each of the securities. The process terminates and there is no need to proceed to any stage 4. The process generally will terminate when the sensitivity report for D shows that the upper limit of the objective function coefficient of D is + infinity or the lower limit is – infinity. In this example the process was halted in stage 2 because negative returns were deemed to be unacceptable.

CONCLUSION

This work shows how to circumvent a problem in a linear programming formulation of the portfolio selection problem. After the optimal portfolio has been established, attention might turn to the question of how the portfolio structure would be altered if any of the expected return parameters of the objective function change. If just one coefficient changes, the standard objective function sensitivity analysis reveals the largest positive or negative change in the expected return on that security that will still result in the optimal portfolio allocations to remain the same. However, the expected returns of the several securities under consideration are quite likely to vary simultaneously and synchronously in response to changing market conditions. It has been shown here how to determine portfolio changes as the expected returns from the several securities under consideration fluctuate. The analysis is limited to the case where all of the fluctuations in expected returns maintain the same ratio to the changes in one of the expected returns. The complete work with references and tables is available from the author.

INVESTIGATION ON TAIWANESE INVESTORS' ASSET PREFERENCES

Angela H-L. Chen, Nanya Institute of Technology, NO. 414, Sec. 3, Jhongshan E. Rd., Jhongli City, Taoyuan County 32091, Taiwan, +886-3- 436-1070, achen@nanya.edu.tw
Zu-Hsu Lee, Marist College, 3399 North Rd., Poughkeepsie, NY 12601, USA, +1-845-575-3000, zuhsu.lee@marist.edu
Beate Klingenberg, Marist College, 3399 North Rd., Poughkeepsie, NY 12601, USA, +1-845-575-3000, beate.klingenberg@marist.edu

ABSTRACT

Investment decisions involve strategic, informed choices. Taiwanese investors have exhibited certain characteristics that deviate from the rationale behind typical finance theories. The AHP (Analytical Hierarchy Process) method is employed to prioritize assets based on investors' preferences in these alternatives when different factors being considered such as market environment, investment amount, expected return rate, risk tolerance and investment type. The result shows that overall, Taiwanese investors prefer stocks to the other assets. When market environment and risk tolerance were considered, mutual funds would be chosen prior to stocks. Whichever criterion is used, bonds turned out to be the least favor asset.
Keywords: asset allocation, portfolio management, analytic hierarch process, behavioral finance, Taiwan

INTRODUCTION

Typical finance models address issues in financial markets based on the efficient-market hypothesis and rationale expectations assumption. Behavioral finance is one subject in the field that deals with investors and their ways of gathering and using information (Fromlet, 2001). How investors learn, adapt and evolve to inter-react with the investment environment (Lovric *et al*., 2008) underlies the plausible assumption of the *Homo economicus* model (Backer, 1976; Rosenberg, 1979). However, although investors may be capable of searching market information efficiently and then determining investment alternatives in an analytical and rational manner, their personal traits (e.g., personality, cognition, emotion) and human and social factors can have a major impact on the final decision making. For example, the investors who have had recent capital gains tend to be more risk-resistant or aggressive than those who just had losses, since the level of risk tolerance of the investor can be dependent on the recent portfolio performance (Langevoort, 2003).

In Taiwan's financial market, individual investors make up the majority trading volume in the Taiwan Stock Exchange (TSE) which exhibits numerous phenomena which thus far cannot be synthesized and fully understood. For example, the market has abnormal trade frequency and each trade generates lower return than that in other part of Asia (Barber *et al*., 2006). It has been a question of whether this can be simply explained by the fact that the less informed, poor trained individuals (Arbel and Strebel, 1983) often trade more frequently, speculatively and suffer lower return than that of smarter and skilled institutional investors.

To study further, this investigation was conducted at dawn of the election for the Taiwanese 12th-term President, which was also the time economic negativities such as the rise of oil price, inflation, and the rising unemployment rate had started to draw worries for Taiwanese investors. Global economic slowdown brought economic challenges and uncertainties to the Taiwanese financial market (e.g., the island's export-oriented economy has left it susceptible to the economic downturn resulting from the current global financial crisis). However, the political campaign resulted in another impact on the market. The Kuomintang (KMT – National Party) had promised a closer economic cooperation agreement with China, and would enlarge the domestic demand by investing in various domestic infrastructure projects. That year, Taiwan's economic performance finished with a 5.7 percent increase in real GDP. Immediately after the election's victory of KMT on March 22, 2008, the financial market jumped more than 6 percent at the opening and closed up nearly 4 percent. Also, Taiwan's currency rose more than 1 percent against the USD, the strongest increase since 1998.

In a volatile market, how Taiwanese investors develop their investment strategy is a question. Whether their preferred portfolios are a rational choice and/or what may be the rationale behind their investment decision are another question. Our discussion centers on the suitability of asset allocation, where asset allocation by definition refers to the process of balancing asset weights in a portfolio within the constraints of an investor's capital resources and investment time horizon, in order to attain the most favorable risk-return tradeoff. A suitable portfolio is a set of assets held that are appropriate to the investor's investment objective(s), financial needs, and level of sophistication. Investors are often advised that their investment portfolio be diversified to minimize the risk, but little attention is paid to other issues that investors may concern.

In this paper, a survey is used to obtain the information of investors' investment preferences and to understand how the factors/criteria such as market environment, investment amount, expected return rate, risk tolerance and investment type affect their selection of portfolio assets. Four investments, stocks, mutual funds, foreign currencies and bonds, representing a wide range of distinct asset classes, are considered when designing the questionnaire. Then, we apply the Analytical Hierarchy Process (AHP) method to analyze the interdependence between the decision factors/criteria and the investors' preferences in assets. Under different criteria, the weights for the assets are calculated by AHP to indicate the priority of each asset to investors. Finally, we discuss the implications behind our results.

METHODOLOGY

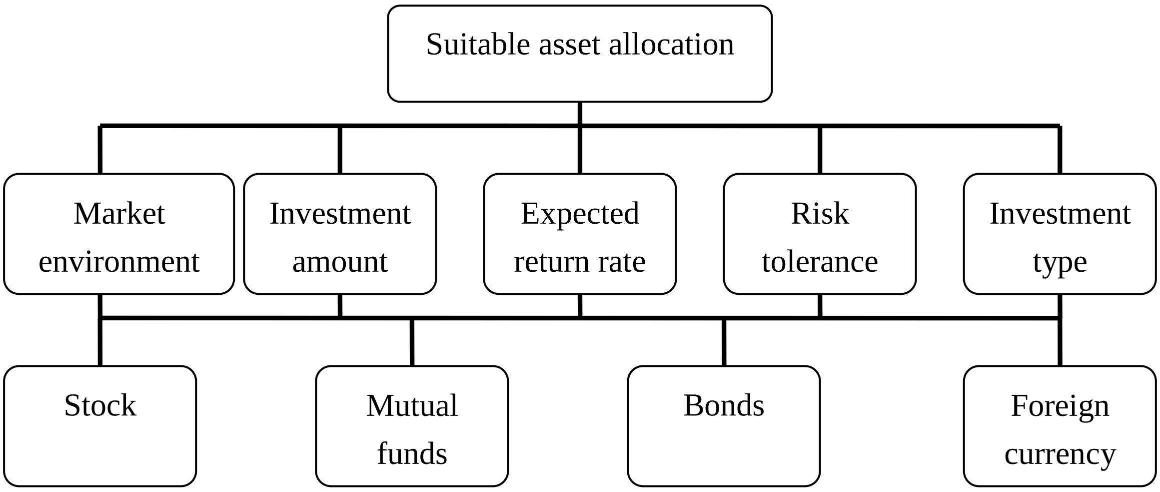
We use AHP (Saaty, 1980) to address the suitability of the portfolio based on an individual's preference. AHP has been a commonly used method to analyze investors' decision factors and investment preferences, in addition to other methods such as Fussy Sets Theory and Delphi Method. However, to our knowledge, no AHP applications have broadly included different types of investment alternatives along with the investor's subjective valuation of assets for the asset allocation problem.

In our AHP procedure, under each investment criterion, different rankings of investment alternatives (i.e., four assets – stocks, mutual funds, foreign currencies, and bonds in our problem) are obtained based on the sample investors' pair-wise comparisons for one alternative to the other using the 9-point rating scale (Saaty, 1980 & 1986). These rankings are expressed in terms of weights and these weights are normalized. In the end, the overall weights for investment alternatives are calculated, where the weight for an alternative may be regarded as the fraction of capital that the investor allocates to this alternative (according to his or her preferences). We use the software of Expert Choice to facilitate AHP calculations and integrate the data from the sample investors.

ANALYSIS OF SURVEY DATA

Market environment, investment amount, expected return rate, risk tolerance and investment type are common considerations (or criteria) addressed in the existing literature and business reports for asset allocation. Market environment refers to all of the forces that affect an investor's decision making in selecting suitable assets. Investment amount means the amount of capital an investor is willing to allocate to an asset. Expeted return rate is the ratio of the gain or loss to the captial that is expected from an asset. Risk tolerance indicates the ability of an investor to cope with the fluctuations in the value of the asset. As for investment type, this refers to the investment method and the investment requirements/ procedure that an investor feels comfortable with. Further, stocks, mutual funds, bonds and foreign currency are four major types of assets that Taiwanese investors choose as investment alternatives in the financial market. Thus, based on the above, we construct a hierarchy for our problem as follows in Figure 1.

FIGURE 1 Decision Hierarchy for Multi-Attribute Asset Allocation



Investors are randomly selected to answer our survey questionnaire and fifty valid responses are collected. Each investor is asked to make comparisons between criteria, and between alternatives under each criterion following the standard AHP procedure. The results are shown in Table 1.

TABLE 1 Rankings under Each Criterion and the Overall for Investment Alternatives

	Market environment	**Investment amount**	**Expected return rate**	**Risk tolerance**	**Investment type**	**Overall priority vector**	**Overall ranking**
Stock (ranking)	0.264 (2)	0.302 (1)	0.299 (1)	0.246 (3)	0.289 (1)	0.2778	**1**
Mutual funds (ranking)	0.310 (1)	0.257 (3)	0.247 (3)	0.278 (1)	0.239 (3)	0.2689	**2**
Bonds (ranking)	0.188 (4)	0.182 (4)	0.193 (4)	0.225 (4)	0.201 (4)	0.1986	**4**
Foreign currency (ranking)	0.238 (3)	0.259 (2)	0.261 (2)	0.251 (2)	0.271 (2)	0.2547	**3**

Table 1 shows that investors concerned with the market environment choose mutual funds first, followed by foreign currency, stock and bonds. However, considering the allocation of capital over assets, they intend to invest the most amount in stocks, and then in foreign currency, mutual funds and bonds in respective order. As far as investors' expectation on the return goes, stock appears to be the most popular one, then foreign currency, mutual funds and bonds. As for whether an asset may meet the degree of uncertainties in the market that an investor is willing to take, Taiwanese investors first consider mutual funds, then foreign currency, stock and bonds. Considering the criterion of investment type, the preference list would be in the order of stock, foreign currency, mutual funds and bonds. Taking into account all of these investment criteria, Taiwan investors generally like stock the most, then mutual funds, foreign currency and bonds.

IMPLICATIONS AND CONCLUSION

In Taiwan's financial market, there are phenomena that cannot be synthesized and fully understood. For example, compared to the trading in other Asian markets, we have seen abnormal trading frequency and each trade usually generates a lower return (Barber *et al*., 2006).

This investigation was conducted at dawn of the election for the Taiwanese 12th-term President. Political campaign and potential economy downturn resulting from the global financial crisis have added unknown variables to the market. Our survey uncovered some interesting behaviors exhibited by Taiwanese investors when facing such uncertainty in the investment environment. The result shows that the order of preference for most investors in four assets is stock, mutual funds, foreign currency and bonds, where stock was clearly the investors' first choice as it was perceived as with highest return and investors were willing to bet with most amounts. In consistency with numerous prior studies' results (e.g., Shu *et al*., 2004; Barber *et al*., 2006; Chang, 2008), Taiwanese investors not only exhibit disposition effect, house money effect, but also a strong risk-seeking behavior. The answer to whether or not such a portfolio of assets in the above preference order at a time of great uncertainty is rational may provide insightful implications behind Taiwanese investors' behavior.

As KMT promised to enlarge the domestic demand by investing in domestic infrastructure projects (the estimated amount was USD $130 billion), a portfolio should be constructed to counterattack the surging price due to the pressure of inflation caused by the tentative, promised massive government spending. Commodities may be flourished under this situation. A less informed, poor trained individual may want to consider a fund which manages industrial or agricultural commodities. In addition, bonds should perform better than stocks as rising interest rates hurt corporate profits and make economic growth difficult. Even cash or foreign currency is capable of providing better return than stocks since governments (or foreign governments) may raise short-term interest rates to give investors incentive to capture the escalating rates through banks' short-term deposit.

Studies show that investments react negatively to uncertainty under certain assumptions (Nakamura, 1999; Saltari and Ticchi, 2005; Femminis, 2006). Should investors predict the deflation would happen after KMT took over the administration (e.g., negative housing market, higher unemployment rate and reductions of government spending), a rational investment strategy may follow the order of bonds, currency, mutual funds and stocks. Bonds should have heavier percentage in the portfolio as the steady income from long-term treasury bonds, for example, would worth more than falling consumer prices. Also, having cash in a bank would be a good risk aversion behavior.

However, the result indicates an overall tendency to choosing stocks as an investment tool for Taiwanese investors. This can be understood by the familiarity breeds investment theory (Huberman, 2001) which asserts that investors tend to invest in the familiar while often ignore the principles of portfolio. Stocks are visible and exposed favorably in the media (Lovric *et al*., 2008); this surely is the case in Taiwan. Further, overconfidence (Odean, 1999) ruled the investors' behavior. Shiller (2000, p.144) believes that most investors "tend to make judgments in uncertain situations by looking for familiar pattern assuming that future patterns will resemble past ones…". Taiwanese investors tend to be more optimistic (Hsu and Shiu, 2007) and thus, trade too often and trade too overconfidently. Mutual funds are managed professionally; however, Taiwanese investors believe their trading success attribute to their own ability. As such, mutual funds came after stocks. The above overconfidence also implies the tendency of aggression in investment rather than conservative alternatives; thus, foreign currency slips to the place after stocks and mutual funds and bonds turn out to be the least favor one.

In sum, this study confirms some behaviors exhibited by Taiwanese investors that deviate from the rationale behind investment decisions assumed in typical finance models. It also shows the role of stocks in the Taiwanese investors' selection of suitable assets. In future research, we intend to explore some measures of the performance of a suitable portfolio based on the preference order of the investor in assets. This may help investment managers develop a good and suitable investment strategy and better serve their clients.

Note: References available upon request from Professor Angela Chen

MEASURING TRANSACTION EXPOSURE USING VALUE AT RISK: AN APPLICATION OF A FIVE-CURRENCY PORTFOLIO

Kashi Khazeh, Perdue School of Business, Salisbury University, 1101 Camden Avenue, Salisbury, Maryland, 21801, email: kxkhazeh@salisbury.edu, phone: 410-543-6328

Robert C. Winder, Luter School of Business, Christopher Newport University, 1 University Place, Newport News, Virginia, 23606, email: rwinder@cnu.edu, phone: 757-594-7004

Fatollah Salimian, Perdue School of Business, Salisbury University, 1101 Camden Avenue, Salisbury, Maryland, 21801, email: fxsalimian@salisbury.edu, phone: 410-543-6321

ABSTRACT

The delta normal approach to "value at risk" is employed for a recent time period to assess the relationship between the transaction risk faced by a hypothetical MNC and the number of currencies to which it is exposed. Consistent with accepted portfolio theory, the maximum downside risk diminishes as the number of currency exposures increases. The rate at which transaction risk falls depends, ultimately, on the standard deviations of the individual exchange rates as well as their correlation coefficients. This paper should provide insight to MNCs as to how to measure, and how to manage, the risk of doing business in foreign currencies.

Keywords: International Finance, Globalization, Diversification Strategies.

INTRODUCTION AND OVERVIEW

This research paper employs the "value at risk" approach to measuring transaction exposure for a hypothetical MNC transacting business in five specific foreign currencies for a recent time period. The key question explored is: to what extent value at risk is reduced by increasing the number of currencies (i.e., from one foreign currency through five foreign currencies). The currencies included in this study are the Swiss franc, the British pound, the Euro, the Canadian dollar, and the Japanese yen (i.e., all with respect to the U.S. dollar). The time period which forms the basis of this study is February 12, 2007 to March 23, 2007. This time period includes thirty consecutive daily observations on the relevant spot exchange rates.

Multinational corporations are exposed to exchange rate risk on an ongoing basis. One form of exchange rate risk is transaction exposure. This is the risk that the MNC's cash flows will be affected by exchange rate changes. Both receivables and payables denominated in foreign currencies add to this risk. "Value at risk" is a probabilistic approach to measuring downside risk (i.e., the maximum loss) that is likely to occur within a specific time frame at a particular level of confidence.

A firm may utilize this method to assess the transaction risk associated with its net cash flows denominated in a specific foreign currency. In this case, the downside risk (i.e., the maximum loss) is a function of the standard deviation in the percentage changes of the particular exchange rate, the (dollar) value of the net cash flow itself, and the desired confidence level. The "value at risk" is positively associated with each of these three variables.

More importantly, an MNC may utilize this basic approach to assess the riskiness of the net cash flows associated with the variety (or "portfolio') of currencies in which it transacts business. Based on standard portfolio theory, the transaction risk (i.e., the maximum loss) is a function of the proportions of the total portfolio in each currency, the standard deviations of the percentage changes in each exchange rate, the correlation coefficients of the percentage changes of the relevant exchange rates, the (dollar) value of the net cash flows, and the confidence level.

Distinctively, a portfolio of currencies whose values are highly volatile vis-à-vis the U.S. dollar (i.e., the standard deviations in percentages changes in the dollar exchange rates are high) will have a high level of transaction risk, ceteris paribus. Portfolios of currencies that possess positive high correlation coefficients will also face more "value at risk," other things equal. On the other hand, portfolios of currencies that have low (or even negative) correlation coefficients will have less value at risk due to internal (or natural) diversification effects.

RELATED LITERATURE

Linsmeier and Pearson [11] provide an excellent overview of the advantages and disadvantages of the most popular approaches to estimating value at risk. Al Janabi [1] provides an excellent primer on the delta normal method. The following articles provide good overviews of VAR methodology or empirical tests: Carrada-Bravo, Hosseini, and Fernandez [4], Tardivo [14], Angelidis and Degiannakis [2], and Chong [5]. The article by Kimball [10] provides an excellent perspective as to why corporations may be prone to miscalculate risk. Platt [13] provides an excellent discussion of the increased use of value at risk resulting from globalization.

While the value at risk methodology is widely employed, there are potential shortcomings. One of these shortcomings is the possibility that the assumption the variable (or variables) in question is normally distributed is incorrect. Articles that explore the implications of nonnormal distributions, including fat tails and how to employ VAR in these cases include Bekiros and Georgoutsos [3], Novak, Dalla, and Giraitis [12], Ferreira [6], and Kaut, Vladimirou, Wallace, and Zenios [7].

This article is an extension of the authors' previous work. See Khazeh and Winder [8] and Khazeh and Winder [9].

EMPIRICAL FINDINGS AND DISCUSSION

Table 1, below, indicates the correlation coefficients between the percentage changes in the five exchange rates for the time period being evaluated.

TABLE 1

	CAD	CHF	GBP	EUR	JPY
CAD	1				
CHF	-0.3219	1			
GBP	0.789634	-0.04265	1		
EUR	0.228513	0.771996	0.378116	1	
JPY	-0.19294	0.448259	0.043088	0.27463	1

Based on the correlation coefficients (above) and the standard deviations of the percentage changes, Table 2 (below) indicates the maximum one-day loss for multiple currency exposures (i.e., as a percent of the MNC's net cash flows denominated in these currencies) at the 95-percent confidence level. As indicated, the foreign currencies (i.e., vis-à-vis the U.S. dollar) were added in the following order: the Canadian dollar, the Swiss franc, the euro, the British pound, and the Japanese yen. In other words, the two-currency portfolio (below) includes the Canadian dollar and the Swiss Franc, the three-currency portfolio includes the Canadian dollar, the Swiss franc, and the euro, etc.

As can be seen from Table 2, value at risk is reduced as the number of currencies in the portfolio expands. One can also observe that the reduction in risk that results from additional currencies diminishes as the number of currencies in the portfolio expands.

TABLE 2

CAD	CAD & CHF	CAD & CHF & EUR
-0.69615	-0.49248	-0.46629
-0.62630	-0.42079	-0.39815
-0.62293	-0.40554	-0.36803
-0.61235	-0.43100	-0.39131
-0.68016	-0.43135	-0.38870
-0.68616	-0.48272	-0.44532
-0.65711	-0.43556	-0.41292
-0.49301	-0.40838	-0.40186
-0.49313	-0.41187	-0.40230
-0.48269	-0.43537	-0.42811
-0.48524	-0.49349	-0.46507
-0.79941	-0.63274	-0.54998
-0.80118	-0.62625	-0.55311
-0.79918	-0.61714	-0.54607
Average	Average	Average
-0.63821	-0.48033	-0.44409

CAD & CHF & EUR & GBP	CAD & CHF & EUR & GBP & JPY
-0.39945	-0.35954
-0.35746	-0.34926
-0.34032	-0.34760
-0.35397	-0.38105
-0.35448	-0.38186
-0.39677	-0.38701
-0.38748	-0.37013
-0.38471	-0.36938
-0.38516	-0.36695
-0.40584	-0.36081
-0.42103	-0.35395
-0.52282	-0.45285
-0.52712	-0.47956
-0.52328	-0.47672

Average	Average
-0.41142	-0.38833

The above computations were made based on a moving 15-day average of the standard deviations for each currency. This approach insures that the most recent information is used to determine the maximum likely loss. In addition, it is assumed in each case that the net cash flows denominated in the various currencies are equally weighted (i.e., in the two-currency portfolio including the Canadian dollar and the Swiss franc, it is assumed that 50 percent of the net cash flows are denominated in Canadian dollars and the other 50 percent are denominated in Swiss francs).

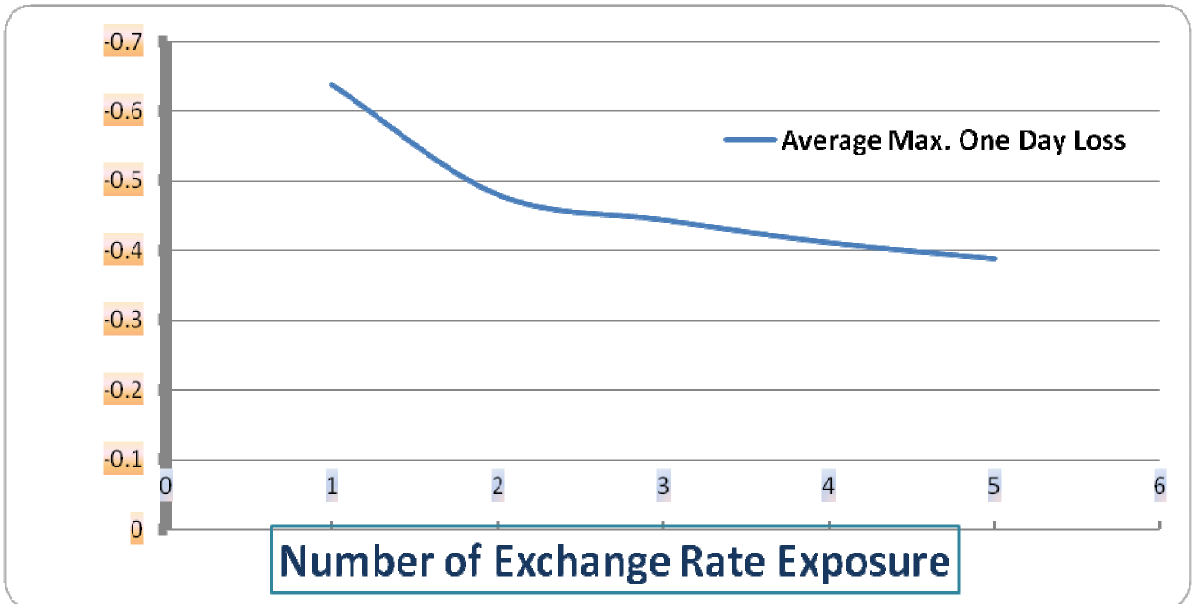
Table 3, below, summarizes the key findings relating to value at risk (i.e., the maximum likely loss) and the number of foreign currencies included in the portfolio.

TABLE 3

Number of Currencies in a Portfolio	Average Maximum One-Day Loss of Portfolio
1	-0.638213339
2	-0.480333794
3	-0.444088183
4	-0.41142049
5	-0.388333251

From another perspective, Figure 1, below, indicates the average maximum one-day loss for multiple currency exposure (i.e., as a percent of the MNC's net cash flows denominated in the foreign currencies)

FIGURE 1



The results, shown above, reveal the key relationships that form the foundation of the value at risk calculations as well the relevant portfolio theory. It is worth pointing out, however, that the rate at which transaction exposure is mitigated as the number of currencies increases will depend on the particular combinations of currencies. The correlation coefficients between the exchange rates seem to differ sufficiently to have a meaningful impact on the value at risk calculations. Because the correlation coefficients between some currencies are positive and high (but less than one), the reduction in value at risk as a result of adding these currencies to the portfolio will be relatively modest. By contrast, because the correlation coefficients between other currencies are negative (e.g., between the Swiss Franc and the Canadian dollar), the reduction in risk resulting from the addition of the latter currencies will be greater.

FUTURE RESEARCH

Future research could include additional combinations of two, three, and four-currency portfolios to determine how various currency combinations affect value at risk. Additional time periods could also be evaluated to assess the stability of the relationships over time. These results should provide key insight and guidance for MNCs in their decisions for managing transaction risk.

REFERENCES

[1] Al Janabi, M. A. M. "Foreign Exchange Trading Risk Management with Value at Risk: Case Analysis of the Moroccan Market." *The Journal of Risk Finance*, 2006, 7(3), 273-291.

[2] Angelidis, T. and Degiannakis, S. "Modelling Risk for Long and Short Trading Positions." *The Journal of Risk Finance*, 2005, 6(3), 226-238.

[3] Bekiros, S. D. and Georgoutsos, D. A. "Extreme Returns and the Contagion Effect between the Foreign Exchange and Stock Market: Evidence from Cyprus." *Applied Financial Economics*, 2008, 18(3), 239-254.

[4] Carrada-Bravo, F., Hosseini, H. K. and Fernandez, L. "Currency Risk Management: Simulating the Canadian Dollar." *International Journal of Managerial Finance*, 2006, 2(1), 77-89.

[5] Chong, J. "Value at Risk from Econometric Models and Implied from Currency Options." *Journal of Forecasting*, 2004, 23(8), 603-620.

[6] Ferreira, M. "Evaluating Interest Rate Covariance Models within a Value at Risk Framework." Journal of Financial Econometrics, 2005, 3(1), 126-168.

[7] Kaut, M., Vladimirou, H., Wallace, W. and Zenios, S. "Stability Analysis of Portfolio Management with Conditional Value at Risk." *Quantitative Finance*, 2007, 7(4), 397-409.

[8] Khazeh, K., and Winder, R., "Measuring Transaction Exposure Using Value at Risk: A Comparison of Two Different Time Periods," presented at the Annual Meeting of the Association for Global Business, Newport Beach, California, November 20-22, 2008.

[9] Khazeh, K. and Winder, R., "Transaction Exposure and Value at Risk: A Practical Application for MNCs," presented at the Annual Meeting of the Southeast Institute for Operations Research and Management Sciences, Myrtle Beach, South Carolina, October 3-4, 2008.

[10] Kimball, R., "Failures in Risk Management," *New England Economic Review*, 2000, Jan/Feb, 3-12.

[11] Linsmeier, T., and Pearson, N., "Value at Risk," *Financial Analysts Journal*, 2000, 56(2), 47-67.

[12] Novak, S. Y., Dalla, V. and Giraitis, L. "Evaluating Currency Risk in Emerging Markets." *Acta Applicandae Mathematicae*, 2007, 97(1), 163-173.

[13] Platt, G. "Increased Volatility in Exchange Rates Spurs Adoption of Best Practices in Currency-Risk Management." *Global Finance*, 2007, 21(5), 72-75.

[14] Tardivo, G. "Value at Risk: the New Benchmark for Managing Market Risk." *Journal of Financial Management and Analysis*, 2002, 15(1), 16-26.

COMPARISON OF RETIREMENT BENEFIT PLANS IN A CHANGING ECONOMY

James Bishop, Bryant University, 1150 Douglas Pike Smithfield RI 02917, (401) 232-6356, jabishop@bryant.edu
Phyllis Schumacher, Bryant University, 1150 Douglas Pike Smithfield RI 02917, (401) 232-6319, pschumac@bryant.edu
Katie Heeder, Bryant University, 1150 Douglas Pike Smithfield RI 02917, (401) 232-6356, kheeder@bryant.edu

ABSTRACT

The focus of this paper is the effect of current market instability on employer sponsored retirement benefits. Today's retirement benefits consist mainly of three types of plans: defined benefit, defined contribution, and "hybrid" plans. Typical benefits provided to the employee by each plan are analyzed in this paper, based on various ages at retirement and lengths of company service. Recent market instability raises questions about an employee's financial security under defined contribution plans and certain hybrid plans. The effect of interest rate and market volatility on each of the different plans is analyzed and recommendations are presented within.

BACKGROUND

Up until thirty years ago, defined benefit (DB) plans were the norm in middle and large sized companies. The benefit was typically a percentage (based on years of service) of an employee's average pay over some period (usually the final five years) of employment. The concept was to support a retiree over their remaining lifetime with approximately 40% - 70% of their final salary while employed. Combining this with Social Security benefits, it was hoped that this benefit would serve a retired employee's economic needs over their remaining lifetime. The cost of providing benefits in DB plans was usually absorbed by the employer, although some plans required participant contributions in order to partially offset the company's expense.

Benefits accrued in DB plans are virtually risk-free. These benefits are defined as a guaranteed annuity payable for the life of the retiree beginning at his/her normal retirement age (usually age 65). If, for some reason, the employer is unable to pay out their benefit obligations, the Pension Benefit Guaranty Corporation (PBGC) insures these benefits up to a maximum monthly amount [9].

Over time, defined benefit plans became ill-suited for employers and employees. For employers, DB plans are expensive to maintain because they often require the use of an actuary and several lawyers to ensure proper compliance with the laws governing DB plans [5]. For employees, the nature of the workplace had changed and long careers at a single company became less common. Under this type of plan, a vested employee (an employee that is accruing benefits), who changes employers would continue to be tied to the company until he/she begins collecting the benefit upon retirement. To meet the needs of a new, mobile workforce, it seemed more desirable for these benefits to be portable so that employees could keep some control over their retirement income when changing companies. As a result, defined contribution (DC) plans became quite commonplace [8].

The most common DC plans, known as 401(k) plans, are defined as individual account balances in which the employee and/or employer each contribute a percentage of the employee's monthly pay into an array of market funds, comprised of mutual funds and bonds, chosen by the employee. At retirement, the accrued account balance is the employee's retirement benefit. Thus, unlike DB plans, DC plans are vulnerable to market instability. Another negative aspect of DC plans is that benefits accrue as a lump sum balance instead of a lifetime annuity as is provided by DB plans. Upon retirement, responsibility is placed on the employee to allocate their lump sum balance into income that will last until death. Due to inadequate planning, many retirees have exhausted their funds and outlived the money they had saved. For the purpose of this study, 401(k) plans are the only type of DC plan that will be considered.

At the same time that defined benefit plans were giving way to defined contribution plans, "hybrid" plans were introduced. Hybrid plans are currently classified under the category of a defined benefit plan even though the intent of the plan is to combine the stability of a DB plan with the portability and lower cost of a DC plan. The most common hybrid plan, known as a cash balance plan, differs from a DC plan in that the contributions made into an account often yield an indexed or fixed interest rate (or a narrow range of rates) chosen by the employer, as opposed to the employee chosen, volatile mutual fund yield of a DC plan. Another major difference between a cash balance plan and a DC plan is that the final account balance has to be offered as a life annuity, just as all other DB plans require.

Another common hybrid plan, known as a Pension Equity Plan (PEP), is computed similarly to the cash balance plan except that the lump sum balance is not payable until the plan's normal retirement age (usually age 65). Whereas cash balance plans provide an employee with a hypothetical investment account payable today, the investment account balance in a PEP plan is a projected value. This value would need to be discounted in order to determine its equivalent value today [2].

Market instability in 2001 and again in 2008 caused employees to lose a significant amount of their retirement savings and has thus raised serious questions about the financial stability of DC plans and certain hybrid plans [1]. The primary objectives of this paper are to compare current retirement plans from an employee's perspective and to consider the ramifications of volatile market conditions on accrued benefits. As such, it is important to note that the benefits calculated for this study are exclusive of participant contributions; therefore, only employer funded benefits are calculated.

RETIREMENT PLAN COMPARISON

In order to compare a typical defined contribution, defined benefit, and hybrid plan benefit for an employee, adjustments need to be made since these benefits are paid in different forms. All three plans will be compared as lump sum values, as well as, life annuities expressed as a percentage of the final pay while employed (salary replacement ratio). A defined benefit calculation is computed as a life annuity and must therefore be converted to a lump sum. Conversely, defined contribution and hybrid calculations are expressed as a lump sum and must be converted to life annuities. The following assumptions are made:

- Current 55 year old with an annual salary of $100,000
- 3% annual salary increase in the future and 3% decrease in the past
- Benefit accrual begins on the date of hire (also called service start age)
- Service start ages of 25, 35, 45, and 55
- Contributory and hybrid plan contributions are made at the end of each year
- Annual interest compounding
- Actuarial conversion factors for various forms and timing of benefit payments are based on current Pension Protection Act assumptions (2010 mortality and transitional yield rates) [4].

Consider three retirement ages: age 55 is a typical early retirement age, age 62 is the first age at which a worker may begin social security payments, and age 65 is a typical normal retirement age as defined in most retirement plans. Finally, as a fair basis of comparison, let us only consider employer funded benefits, exclusive of participant contributions.

First, we compute a defined contribution plan benefit assuming a 5% of pay employer contribution and a market rate of return of 7.5%. The 5% contribution assumption is based on an approximate average large company DC plan (range is roughly 2% - 8%). A market yield of 7.5% is about average over the past 60 years based on the S&P index. Contributions will be considered to be made at the end of each year into an account earning 7.5% compound interest until retirement. The resulting lump sum retirement values and salary replacement ratio's (annual life annuities divided by final salary at retirement) are shown in the table below:

DC Plan - Lump Sum Values Payable at Retirement Age Amounts are converted to an annual life annuity and expressed as a percent of final salary at retirement)				
	Service Start Age			
Age at Retirement	**25**	**35**	**45**	**55**
55	$ 289,658.40 (20%)	$ 150,214.68 (10%)	$ 59,289.00 (4%)	$ -
62	$ 528,243.66 (33%)	$ 296,899.68 (19%)	$ 146,049.51 (9%)	$ 47,686.14
65	$ 676,674.57 (42%)	$ 389,276.67 (24%)	$ 201,875.98 (12%)	$ 79,679.46

Next, let's consider a typical final five year salary Defined Benefit plan with the following formula for an annual benefit paid starting at retirement age for the life of the employee:

Annual benefit = 1.5% × FAP5 × BS; where FAP5 represents the average of the employee's final five full years of pay at the company, and, BS represents the number of years the employee worked for the company (benefit service). The accrual percentage of 1.5% is chosen as the middle of a typical range for DB plans (typically 1% - 2%) [3].

In order to compare the annual annuity payable in the Defined Benefit plan with the lump sum payable in the Defined Contribution plan, we must convert the DB plan annuity to a lump sum using actuarial equivalence factors. This conversion is dependent on the interest and mortality assumptions at the time of retirement. Using reasonable current assumptions (year 2010 factors under the Pension Protection Act), the resulting table of lump sum values is below:

DB Plan - Lump Sum Values Payable at Retirement Age (converted to a life annuity and expressed as a percent of final salary at retirement)				
	Service Start Age			
Age at Retirement	**25**	**35**	**45**	**55**
55	$ 280,116.09 (19%)	$ 186,744.06 (13%)	$ 93,372.03 (6%)	$ - (0%)
62	$ 632,680.49 (40%)	$ 461,685.77 (29%)	$ 290,691.04 (18%)	$ 119,696.31 (8%)
65	$ 892,467.11 (55%)	$ 669,350.34 (41%)	$ 446,233.56 (27%)	$ 223,116.78 (14%)

Note that long careers favor the defined benefit plans. This is partly due to the fact that DB plans are actuarially reduced for early commencement, so an employee that works up until 62 or 65 does not get a stiff reduction. DC and DB plans give very similar benefits at age 55. It is also noteworthy that many DB plans give an extra incentive for early retirement (a modest reduction rather than an actuarial reduction) and so these plans would compare favorably even at age 55.

Finally, let's examine a hybrid plan. Consider a Pension Equity Plan (PEP) with an annual contribution of 9% of pay. Once again, a moderate value has been chosen from a range of approximately 6% - 12% for such plan types [7]. The interest accumulation rate is 3.5% (a typical fixed guaranteed rate, not a market assumption). One oddity inherent in this type of retirement plan is that the balance is always assumed payable at the plan's normal retirement age. If the employee quits/retires prior to that time, then the interest stops accumulating AND there is a reduction in benefit for taking the money early. The following lump sum table results:

PEP Plan - Lump Sum Values Payable at Retirement Age (converted to a life annuity and expressed as a percent of final salary at retirement)				
	Service Start Age			
Age at Retirement	**25**	**35**	**45**	**55**
55	$ 162,331.67 (11%)	$ 105,580.63 (7%)	$ 51,512.35 (3%)	$ - (0%)
62	$ 363,803.08 (23%)	$ 258,965.14 (16%)	$ 159,083.14 (10%)	$ 63,922.80 (4%)
65	$ 517,043.17 (32%)	$ 378,247.04 (23%)	$ 246,012.13 (15%)	$ 120,028.29 (7%)

The PEP plan does not compare favorably to either DB or DC plans. The impact of these less generous plans will be felt by employees in the years to come.

MULTIPLE RETIREMENT PLANS

Finally, it is worthwhile to examine a very common career path in the modern workplace. Within in the past 20 years, many workers either experience at least one change in employers (accruing benefits under multiple retirement plans), or, have a plan change at their lone place of employment. Here we examine an example of a switch of plan types ten years after the beginning of an employee's career. Below is a table representing a common change from a traditional DB plan to a hybrid plan:

Value of a Life Annuity Expressed as a Percent of Final Salary at Retirement			
Service Start Age (assumes a change of plan 10 years after Service Start Age)			
Age at Retirement	**25**	**35**	**45**
55	23%	18%	14%
62	27%	21%	16%
65	29%	23%	18%

It can be seen that all scenario's lead to retirement with less than 30% of the annual income enjoyed by the individual while actively employed.

STOCK MARKET INVESTMENT

Now let us consider how market fluctuations affect retirement benefit amounts. Most employee's keep their DC balances in the stock market, but for employee's over age 55 it makes sense to shift these funds over to bonds (if they are an available alternative). The S&P stock market index is a stable benchmark for tracking investment return on Defined Contribution accounts. We take a look at moving averages over 5, 10, and 20 years for the annual investment returns of the S&P. The following table results:

Number of Years			
	5	**10**	**20**
Mean Increase	7.8%	7.7%	7.3%
Standard Deviation of Increase	6.9%	4.7%	3.2%
Probability of loss	13.0%	5.3%	1.0%

S&P annual increases from 1950 – 2008

How does market return impact DC balances? The standard deviation of market increases over longer periods (10 and 20 years) is relatively low. Therefore, DC Plans are reasonably safe unless the employee is close to retirement. If an employee happens to be 60 years old and is looking to retire at age 65, then there is approximately a 13% chance that the market return on their DC account balance will be negative based on past market performance. This figure is based on the normal distribution z-score for the 5 year average mean and SD.

CONCLUSION

So what does the future hold for retirement plans? Part of the answer for this depends partly on whether there is a sentiment of protecting people against themselves, in that, account money can automatically shift to a safer investment when close to retirement. Of course, not everyone is averse to the risk associated with the stock market.

Cash balance and hybrid plans are stable and are a good option, they could (and often do) have a floor for the minimum interest earned on the accounts for any given year. In order to make these plan types affordable to the employer, there would have to be a new and simpler set of regulations to match these plan types. Currently the rules for these plans fall under the umbrella of all defined benefit plans.

The future of retirement plans should naturally drift towards a cash balance benefit with a required stable conversion to a life annuity upon retirement. The advantage of the cash balance plan is its steady, intuitive balance increase, combined with the security of a life annuity at retirement. This may require new legislative considerations because the rules governing such plans were originally developed for traditional defined benefit plans.

There is no recommendation that will help the employer afford to support retiree's in all cases (or even many cases), but allowing an employee to plan for their own retirement with some surety of their income is the desired goal.

REFERENCES

[1] Coombes, A. (2009, October 6). *Steady savers gained over 5-year period*. Retrieved from http://www.marketwatch.com/story/401ks-took-big-but-not-devastating-hit-in-2008-2009-10-06.

[2] Green, L. (2003, October 29). *What is a pension equity plan?*. Retrieved from http://www.bls.gov/opub/cwc/cm20031016ar01p1.htm

[3] Hamilton, B, & Burns, B. (2001, December 31). *Reinventing retirement income in america.* Retrieved from http://www.ncpa.org/pub/st248?pg=5

[4] IRS. (2009, September 5). *Choosing a retirement plan: defined benefit plan*. Retrieved from http://www.irs.gov/retirement/article/0,,id=108950,00.html

[5] IRS. (2008). *Updated Static Mortality Tables for the Years 2009 Through 2013.* Internal Revenue Bulletin 2008-42.

[6] Kamenir, J. (2009). *How to Make Defined Benefit Pension Plans Attractive to 21st Century* Employers. Benefits Quarterly, 25(2), 51-56. http://search.ebscohost.com

[7] Lowman, T. (2000, July 7). *Actuarial aspects of cash balance plans*. Retrieved from http://www.soa.org/files/pdf/actuarial_aspects.pdf

[8] McGill, D, Brown, K, Haley, J, & Schieber, S. (2005). *Fundamentals of private pensions*. Oxford, NY: Oxford University Press.

[9] PBGC. (2009, October 27). *PBGC Announces Maximum Insurance Benefit for 2010*. http://www.pbgc.gov/media/news-archive/news-releases/2009/pr10-02.html

DEALING WITH FINANCIAL TRAUMA: A BRIEFING FOR FINANCIAL ADVISORS

Thomas P. Langdon, Roger Williams University, (401) 254-5356, tlangdon@rwu.edu
Barbara S. Poole, Roger Williams University, (401) 254-3182, bpoole@rwu.edu

ABSTRACT

This paper introduces the psychology and treatment of trauma to financial advisors who are coping with both their own and their clients' financial traumas. As a result of this work, financial advisors can be better equipped to identify signs of financial trauma, develop an awareness of the resources available to support individuals undergoing financial trauma, provide educational information and/or referrals for their clients, and consider integrating some level of support in their practices.

Keywords: financial advisor, mental health, therapy, trauma

INTRODUCTION

Eonomic pundits suggest that following a series of massive financial crises, Americans should get accustomed to the "new normal". This new normal is characterized by increasingly conservative financial behavior, including lower household debt, higher personal savings, and reduced consumption (Galston, 2009). Despite President Obama's assurances to the contrary (State of the Union, 2010), households expect higher unemployment and increased taxes. Clearly, the carefree days of irrational exuberance are past, and not expected to return any time soon.

No longer the glamorous industry of romanticized excesses, the American financial services industry now faces angry taxpayers outraged at the tax dollars spent on bailouts, and at the continuing excesses in executive compensation. At this writing, the Obama administration has proposed a fee for financial firms and automobile manufacturers that received bailout funds "in hopes of soothing the public's anger at financial firms" (Solomon and Paletta, 2010). Shareholders in financial firms have already begun to take action in protest. For example, the Security Police and Fire Professionals of America Retirement Fund, a shareholder of Goldman Sachs Group, Inc., has sued the organization for paying record 2009 compensation to executives, charging that the firm prospered not as a result of executive performance, but due to the bailout (PR Newswire, 2009).

In addition to adjusting to the new normal environment, many households are still reeling from the economic blows of the past few years. Formerly dependable industries such as financial services and auto manufacturers have required government bailouts to remain in business. Investors have suffered losses from financial schemes such as those perpetrated at Enron and Madoff Investment Securities. In 2008, a record 3,157,806 foreclosures were filed, an increase of 81% over 2007, and a 225% increase over 2006 (RealtyTrac®, 2009). In a tight economic climate, foreclosures and car repossessions go hand in hand. In 2008, 1.67 million cars and trucks were repossessed, also a record volume, representing a 12% increase over 2007 (RealtyTrak, 2009).

In the current U. S. economy, few are exempt from the direct or indirect effects of the economic downturn. As unemployment reaches heights not seen since the Great Depression, households may adjust their lifestyles, delay retirement, raise their credit balances, or jeopardize their children's educational funding. Business owners face the challenges of managing a business in anticipation of a potentially prolonged sluggish economy. Those who appear otherwise somewhat immune to the economic downturn may contend with friends or relatives seeking loans, jobs, or living quarters.

Financial advisors have not escaped the backlash of frustrated clients. These feelings may be manifested in reduced client retention, hostile meetings, or filings of official complaints. Financial advisors may be well advised to become skilled in managing financially traumatized clients in order to avoid the experience of one German financial advisor. In June 2009 five retirees who had lost significant investment funds were arrested for allegedly kidnapping, beating, and threatening their financial advisor with death (CNN, 2009).

Financial advisors working with financially distressed clients can suffer compassion fatigue, a phenomenon caused by extended periods of emotional involvement with traumatized individuals. The term originated in the medical profession but can apply to a variety of professionals who work with these clients. "Professionals who listen to clients' stories of fear, pain, and suffering may feel similar fear, pain, and suffering because they care. Sometimes we feel we are losing our sense of self to the clients we serve" (Figley, 1995). Financial advisors often lack the training and support to deal with compassion fatigue and can suffer the same type of symptoms as a directly traumatized individual. Medical researchers have identified that these symptoms include feelings of apathy, isolation, moodiness, sleep disturbance, anger, and somatic symptoms (Figley).

Ranked the best job in America in 2001, financial planning had dropped to number 65 by 2009 (Krantz, 2002, 2010). These rankings are based on the following five factors:

- the work environment, both physical and emotional aspects
- income and growth potential
- employment outlook including employment growth, income growth potential, and unemployment
- physical demands
- stress, based on 21 factors including quotas, deadlines, advocacy, competitiveness, precision and initiative required, and meeting the public.

Clearly, the demands on financial advisors have changed.

Financial advisors, while themselves still adjusting to new normal, must guide their clients' efforts to adjust to new realities. The transition to a new normal environment demands more time reassuring clients, redefining goals, and reassessing strategies. These rational conversations, however, may often take place with clients who are feeling less than rational. These clients may enter their financial advising sessions with greater uncertainty about their financial plans, which had previously been considered well thought out and on course.

Financial advisors may need help preparing their clients to discuss their financial situations from a well-reasoned perspective. While this need is not new, it becomes increasingly important in light of the current state of the economy, for the both client's and the advisor's well being. Clients must overcome their shock, confusion, and anger in order to understand their current financial situations and review and revise plans for the future. They may benefit from referrals to psychology professionals versed in helping trauma patients recover from their financial shocks.

The purpose of this paper is to introduce the psychology and treatment of trauma to financial advisors who are coping with their own financial trauma as well as those of their clients. As a result of this work, financial advisors can be better equipped to identify signs of financial trauma in themselves and their clients, develop an awareness of the resources available to support individuals undergoing financial trauma, provide educational information and/or referrals for their clients, and consider integrating some level of support in their practices, especially those who use a holistic life planning approach.

BACKGROUND

In the mental health field, trauma is viewed as an event "that is emotionally painful, distressful, or shocking, which often results in lasting mental and physical effects" (NIMH, 2001). Therapists focus on treating the emotional and physical response to trauma. A particular event can be traumatic to some patients but not to others, as determined by the subsequent symptoms.

Individuals experiencing traumatic events often suffer from anxiety and a heightened awareness of the uncertainties in their lives. Early behavioral finance studies support that individuals are willing to pay a premium in order to reduce uncertainty (Kahnemann, Slovik, and Tversky, 1979). Citing the work of Frieda Fromm-Reichman (1950), Yalom (2005) emphasized the role of uncertainty in the process of producing anxiety by stating, "The awareness that one is not one's own helmsman…that one's perceptions and behavior are controlled by irrational forces, is itself a common fundamental source of anxiety."

Recent events have reminded investors that they have limited control over their financial destinies. When financial plans, which create the illusion of control, go awry, traumatic symptoms may be elicited that exceed those precipitated by natural causes. Psychologists have observed that human made disasters cause survivors more distress and long-term impairment than natural disasters such as floods and tornados (Howard and Goelitz, 2004). As a result, the emotional impact of events for which clients can blame themselves for their own bad judgment, or can blame others such as their advisors, could exacerbate trauma symptoms.

RESOURCES AVAILABLE

Financial advisors dealing with their own or their clients' traumas can be helped by understanding the resources available for identification of the problem. This section provides an introductory discussion of these resources. In order to differentiate clearly between the roles of the financial advisor and the psychological therapist, the following nomenclature has been adopted for this paper. The individual who has experienced financial trauma will be referred to as a trauma client when the financial advisor-client relationship is referenced. This individual will be referred to as the trauma patient when the therapist-patient relationship is referenced.

Mental Health First Aid (MHFA) training

The National Council for Community Behavioral Healthcare (NCCBH) has initiated a pilot 12-hour Mental Health First Aid program designed to help participants "recognize and respond to" mental health distress (Gallant, 2009). Volunteer training is available in limited areas now, with plans to expand nationally. Research on the impact of the original Australian program indicate that the training increases mental health literacy, empowers the volunteer to help individuals in crisis, connects individuals to services, and reduces stigma associated with mental health issues (NCCBH, 2009). Financial advisors may find that a MHFA-trained staff member could be helpful for identifying clients who need therapeutic support in addition to financial advice.

Psychological First Aid (PFA)

Psychological First Aid (PFA) is a short term immediate intervention for survivors of community disasters. In general, the goal of PFA is to "reduce initial distress, and to foster short and long-term adaptive functioning" (National Child Traumatic Stress Network, 2006). Citing Everly and Flynn (2005), Everly, Phillips, Kane, and Feldman (2006) explained that FPA may be intended to achieve any of the following:

- provision of information/education,
- provision of comfort and support,
- acceleration of recovery,
- promotion of mental health, or
- facilitation of access to continued or escalated care.

From a clinical standpoint, when groups of individuals are exposed to the same traumatic event, there is a strong rationale for implementing PFA practices in their natural homogeneous cohort (Everly, Phillips, Kane, and Feldman, 2006). While PFA is typically associated with treating survivors of natural disasters such as the recent Haitian earthquake, the technique can be appropriate for those who share similar financial trauma such as unemployment or investment scams.

Psychoeducation

Psychoeducation is a mental health intervention that circumvents the stigma typically associated with receiving help (Howard and Goelitz, 2004), and is an important component of any response to a community disaster (Serani, 2004; Howard and Goelitz, 2004; Scheidlinger and Kahn, 2005; Haen, 2005). The focus of psychoeducation is on empowering participants with knowledge concerning typical reactions to disaster and common coping skills, while avoiding "overhelping" those who are naturally recovering (Howard and Goelitz). While participating in a community education project, Howard and Goelitz noted that their psychoeducational session participants were relieved to learn that their trauma reactions were normal, and that they felt supported by the other resources made available to them in the sessions. They indicated "Disaster survivors often feel they are ‘going crazy’; psychoeducation help dispel this myth and enable survivors to recognize their experiences as normative reactions to disaster”.

In addition to social support, psychoeducation promotes empowerment and self-efficacy. The philosophy behind psychoeducation respects participants’ competence to make appropriate mental health decisions when properly informed, empowering participants to access resources if necessary (Howard and Goelitz, 2004). As a result, linking psychoeducation presentations with concrete services is recommended for reaching the broadest possible audience (Howard and Goelitz). Psychoeducation services for financial trauma may include consumer awareness and financial literacy resources.

Trauma-focused healing groups

Trauma-focused healing groups, which include support groups, debriefing groups, and crisis intervention groups, are intended to facilitate acceptance of the situation and restore functionality. These short term groups are directed toward quickly achieving specific, measurable goals. They tend to concentrate on reconstruction of the traumatic event with emphasis on the patient’s feelings and misconceptions, followed by development of plans for future coping and adaptation (Scheidlinger and Kahn, 2005).

Support groups, also called self-help groups, provide emotional support for participants with a common problem, and are usually not organized or facilitated by mental health professionals. While traditional support groups meet regularly in a physical location, support groups have proliferated in the online community. For example, Bernie Madoff’s Ponzi scheme spawned several online support groups which provide not only on emotional support but also updates and advocacy (Brookman, 2009).

Debriefing groups employ “a psychological and educational group process that aims to reduce the impact of a traumatic event” (Webb , 2005). Debriefing groups encourage patients to review their personal responses following a recent crisis so that their experience can be “normalized” as appropriate reactions, given the stressful situation (Webb). Similarly, in crisis intervention groups, “After brief introductions and minimal efforts to establish at least a degree of connectedness …, group discussions are focused on the reality of the event (what actually happened), followed by the appropriate expression of feelings of loss and mourning” (Scheidlinger and Kahn, 2005).

Group and individual therapy

If interventions described above are not sufficient to help traumatized individuals, group therapy can be a viable next step. A therapist treating trauma patients focuses on two primary goals: (1) relieving the consequences of trauma, such as hyperarousal, numbing, dissociation, and reliving of the event; and (2) avoiding “chain shock”, which occurs when the patient links together two or more unrelated traumatic events, heightening the patient’s trauma (Serani, 2004). Effective treatment may be helpful for financial planning clients who want to overcome extreme risk aversion resulting from financial trauma and to restore a rational re-assessment of their financial risks.

Frequently, group therapy is more effective than individual therapy (Corey, 2008). Group therapy may be particularly appropriate for use with individuals exposed to large-scale traumatic events. The goal of the therapy is the reconstruction of a sense of security for the patient, and may be helpful for financial planning clients in need of recovering sufficient functionality to begin financial reassessment and planning.

Once the objectives of trauma-focused healing activities have been met, group and individual therapy typically employs classic psychotherapy techniques to focus on personality organization or repair (Scheidlinger and Kahn, 2005). "Psychodynamic therapies…stress the importance of covert ideations and conflicts as being responsible for current feelings and symptoms." (Scheidlinger and Kahn). Classic psychotherapy techniques such as cognitive behavioral therapy and existential therapy tend to require a long term commitment in order to achieve objectives.

Cognitive behavioral therapy uses learning principles to transform the thought processes and behavior of the patient (Melton, Petrila, Poythress, and Slobogin, 2007). The theory underlying cognitive behavioral therapy postulates that psychological distress is the result of impaired information-processing and disruption in patterns of social behavioral reinforcement (Yalom and Leszcz, 2005). Cognitive behavioral therapists "attempt to access and illuminate these thoughts through probing, Socratic questioning, and the encouragement of self-examination and self-monitoring" (Yalom and Leszcz).

Existential therapy is a way of thinking as opposed to a particular style of practicing group therapy (Corey, 2008). Existential group therapy may give trauma patients the opportunity to address and work through many of their feelings elicited by the traumatic event, including thoughts about death, isolation, and meaninglessness. In some respects, the end result of existential group therapy for trauma patients may be to achieve Yalom and Leszcz's (2005) therapeutic factor of universality. Trauma patients who understand the universality of their financial concerns may have taken the first step toward healing, as well as reassessing what is really important in their lives. As a result, a financial planning client may re-think their relationship to money, as well as their financial goals and strategies.

CONCLUSION

Given the current economic environment and the stress of adjusting to the new normal, financial advisors may be challenged to cope with their own financial trauma as well as those of their clients. This paper has introduced financial advisors to the psychology of trauma, its aftereffects, and treatment. With this awareness, financial advisors may consider steps they can take to integrate education, referrals, and client support in their practices.

REFERENCES

References available upon request from either author.

Director Affiliations and Acquisition Announcement Performance

Chia-Wei Chen, Tunghai University, Taichung, Taiwan, Email: achen@thu.edu.tw

J. Barry Lin*, Simmons College, Boston, MA 02215, Email: j.lin@simmons.edu

Abstract

Expanding on earlier findings that outside director characteristics are significant factors on firm performance, this paper focuses on the significance of their leadership role and affiliations. Examining market reaction to acquisitions, we find acquirers with busy outside directors holding a CEO title in S&P 500 firms or holding diversified directorships obtain higher cumulative abnormal returns during the announcement date than firms with other busy outside directors. While multiple directorships could lead to ineffective monitoring, our finding highlights the potential advantage of certain busy outside directors. Large firm leadership position and diverse experiences appear to be associated with significant positive impact.

Keywords: Board of directors, Multiple directorships, Mergers and acquisitions
JEL classification: G39

*corresponding author

1. Introduction

Outside directorships signal the reputation of a valuable director [Fama (1980) and Fama and Jensen (1983)]. Supporting this notion, researchers find evidence that the likelihood for outside directors to obtain addition board seats is related to the performance of the firm in which they serve on the board [see Ferris, Jagannathan, and Pritchard (2003) and Fich and Shivdasani (2006)]. Similarly, Harris and Shimizu (2004) show that multiple directorships could be important sources of knowledge that helps to improve shareholder value during acquisitions. However, other recent studies suggest that outside directors with multiple directorships, also known as busy outside directors, serve less actively on board functions and therefore may shirk their responsibilities in protecting shareholder wealth. Consistent with this hypothesis, such directors are positively associated with, among other problems, the likelihood of corporate financial fraud [Beasley (1996)] and excessive CEO compensation [Core, Holthausen, and Larcker (1999)]. They are also negatively associated with firm performance [Fich and Shivdasani (2006)]. While the debate on the value of busy outside directors focuses on two competing hypotheses, the *reputation* and the *busyness hypotheses* [see Ferris, Jagannathan, and Pritchard (2003)], current research is scant in examining the link between shareholder wealth and the characteristics of individual busy outside directors.

In this study, we provide additional evidence on the value of busy outside directors by examining their influence on wealth gains in mergers and acquisitions. With a sample of 854 acquisitions from 1998 to 2004, we find that busy outside directors are not all the same. Although busy outside directors in general are negatively associated with acquirer returns during the announcement of acquisitions, those with a CEO title in an S&P 500 firm and those with outside directorships in different industries do not have a similar negative association. In addition, our results indicate that multiple directorships do not reduce the value of outside directors who have a CEO title in an S&P 500 firm. Our results, therefore, suggest that the different characteristics of individual busy directors can have different effects on firm value. These findings may help in understanding the seemingly conflicting relationships between firm value and busy outside directors found in previous studies.

Directors from large firms or directors who are CEOs of large firms are more likely to receive additional directorships [Ferris, Jagannathan, and Pritchard (2003)]. The size and complexity of the operations in large firms suggests that these directors may be more skillful and may provide broader networking contacts to firms in which they serve on their boards. Similarly, Fama and Jensen (1983) indicate that serving as directors on other boards could signal the executive's competence. Kaplan and Reishus (1990), Shivdasani (1993), and Brickley, Linck, and Coles (1999) among others offer empirical evidence to support this notion. Fich (2005) further finds a positive shareholder reaction to the appointment of outside directors who are CEOs of other firms. With this in mind, the value of a director's skills and networking contacts may be even more pronounced if the director is also the CEO of an S&P 500 firm, which tends to be a leader in its industry [see Cai (2007)]. Since directors with CEO titles in S&P 500 firms are likely to be attractive candidates for other boards, do these directors lead to better firm performance than do other busy outside directors who are not S&P 500 firm CEOs? Furthermore, do multiple directorships add value to outside directors with CEO titles in S&P 500 firms? Do busy outside directors with a CEO title in a S&P 500 firm differ from busy outside directors with a CEO title in a non-S&P 500 firm?

2. Empirical Results

In Table 5, we apply OLS regression to test the relationship between acquirer returns and busy outside directors. The dependent variable is the five-day cumulative abnormal return [CARs (-2, 2)]. In regression (1), we first test the relationship between busy outside directors and acquirer returns without considering the characteristics of busy outside directors. Consistent with Ahn, Jiraporn, and Kim (2008), we find that acquirer returns are negatively associated with the percentage of busy outside directors, suggesting that multiple directorships reduce the effectiveness of monitoring by outside directors [Fich and Shivdasani (2006)]. However, when variables indicating the characteristics of busy outside directors are added into regressions, we find that some types of busy outside directors are differently associated with acquirer returns. In regression (3), for example, the coefficient for the percentage of non-industry busy outside directors suggests that these busy outside directors do not reduce acquirer returns. While a one percent increase of busy outside directors reduces five-day CARs by about 0.05 percentage points, if the increased busy outside directors are non-industry busy outside directors, the reduced CARs could be recovered and the total impact from non-industry busy outside directors on acquirer returns is not different from non-busy outside directors (i.e. 0.046 percent increase in five-day CARs).

The coefficient for the percentage of SP500-CEO busy outside directors, as shown in regression (4), suggests that busy outside directors with a CEO title in an S&P 500 firm are not associated with negative acquirer returns. One percentage increase of busy outside directors with a CEO title in an S&P 500 firm potentially recovers the negative acquirer returns caused by multiple directorships held by busy outside directors, and in fact increases acquirer returns by about 0.03 percentage points [i.e. 0.08 + (-0.05)] although this positive return is not significantly different from zero. However, similar to non-industry busy outside directors, SP500-CEO busy outside directors are not different from non-busy outside directors and also improve acquirer returns by about 0.04 percentage points.

For industry busy outside directors and non-SP500 busy outside directors, as shown respectively in regression (2) and (5), we do not find positive or significant coefficients for these variables, suggesting that these busy outside directors are not different from most busy outside directors. A one percent increase of industry or non-SP500 busy outside directors could reduce acquirer returns by 0.04 or 0.03 percentage points, respectively. Although the coefficients for the percentage of busy outside directors are different in regression (2) and (5) because of the potential difference between industry and non-SP500 busy outside directors, the difference is not significantly different from zero. Combining the results from all regressions reported in Table 5, we see that while busy outside directors are negatively associated with five-day CARs during acquisitions, not all busy outside directors are the same. Non-industry busy outside directors and SP500-CEO busy outside directors are associated with more outside directorships, directorships in different industries, and directorships in large firms. Potentially, these directors with their great diversity of experience could be valuable source of knowledge to the firms in which they serve on their boards. However, while we find these directors could be better than other busy outside directors, we do not find the evidence that these directors are better than non-busy outside directors.[1]

Our results also indicate a positive association between the percentage of outside directors and acquirer returns during acquisitions. A one percent increase of outside directors, without considering the characteristics of individual outside directors, improves five-day CARs by about 0.05 percentage points.

3. Multiple Directorships and CEO Title

A. Do Multiple Directorships Add Value to Directors Who Are CEOs of Other Firms?

In regression (4) of Table 5, we report a positive and significant coefficient for the percentage of SP500-CEO busy outside directors, indicating that these busy outside directors could offset the negative impact from multiple directorships held by busy outside directors on acquirer returns. However, these busy outside directors with a CEO title in an S&P 500 firm are not better than non-busy outside directors. We might, therefore, question whether multiple directorships reduce the value of holding a CEO title in an S&P 500 firm. To address this issue, we develop a variable, percentage of SP500-CEO non-busy outside directors, to compare with the percentage of SP500-CEO busy outside directors. As shown in Table 8, while the coefficients for the percentage of SP500-CEO non-busy outside directors are insignificant in all regressions, we observe positive and significant coefficients for the percentage of SP500-CEO busy outside directors. The positive impact from SP500-CEO busy outside directors on acquirer returns recovers the negative impact from multiple directorships held by busy outside directors. Thus,

[1] It is possible that busy outside directors with a CEO title in S&P 500 firms or with directorships in other industries are invited to join the board before the announcement of acquisitions to enhance shareholder wealth. However, the average (median) tenure in the sample is 6.77 (5) years for busy outside directors with a CEO title in S&P 500 firms and is 8.19 (6) years for busy outside directors with directorships in other industries. These tenure characteristics suggest that strategically hiring busy outside directors shortly before a merger announcement is not prevalent.

our results indicate that CEOs of S&P 500 firms who hold multiple directorships are not less valuable members of the board than CEOs of S&P 500 firms who are not busy.

B. Are SP500-CEO Busy Outside Directors Better than Non-SP500-CEO Busy Outside Directors?

Finally, if busy outside directors with a CEO title in an S&P 500 firm are better than other busy outside directors, we might question whether a CEO title in an S&P 500 firm is different from a CEO title in a non-S&P 500 firm. To address this issue, we measure the percentage of non-SP500-CEO busy outside directors to compare with the percentage of SP500-CEO busy outside directors. As reported in Table 9, the relatively small and insignificant coefficients for the percentage of non-SP500-CEO busy outside directors indicate that busy outside directors with a CEO title in a non-S&P 500 firm are different from busy outside directors with a CEO title in an S&P 500 firm. While busy outside directors with a CEO title in an S&P 500 firm do not exhibit the negative impact from multiple directorships held by busy outside directors on acquirer returns as we found in previous sections, similar evidence is not found for busy outside directors with a CEO title in a non-S&P 500 firm. It supports the notion that management skills in relatively large, well-known firms add value to outside directors and therefore enhance shareholder wealth.

4. Summary and Conclusion

Busy directors with multiple directorships in different industries and directors with a CEO title in an S&P 500 firm can be valuable sources of knowledge and experience to firms. Therefore, despite their busyness, these directors may still be good candidates for corporate boards. Fich (2005) suggests that some outside directors could be better than the others. Our results indicate that busy outside directors with diverse outside directorships and management skills in large, well-known firms are different from other busy directors and may be valuable to acquiring firms. On a more general level, our empirical evidence suggests that director characteristic, affiliations, and experiences are important factors that as yet have not been fully examined. Our empirical findings are not only important for portfolio investors but also provide useful insights for M&A arbitrageurs who can utilize outside director affiliations as a factor in their trading strategy.

Table 5: Busy Outside Directors and Acquirer Returns

Variable	(1)	(2)	(3)	(4)	(5)
Percentage of busy outside directors	-3.719**	-3.553**	-5.426***	-5.271***	-2.603*
	(-2.38)	(-2.47)	(-2.78)	(-3.01)	(-1.69)
Percentage of industry busy outside directors		-1.439			
		(-0.24)			
Percentage of non-industry busy outside directors			5.726**		
			(2.16)		
Percentage of SP500-CEO busy outside directors				8.034**	
				(2.40)	
Percentage of non-SP500 busy outside directors					-3.267
					(-1.19)
Percentage of outside directors	4.695*	4.596*	4.623*	4.300*	4.659*
	(1.90)	(1.87)	(1.88)	(1.74)	(1.89)

Log (board size)	0.602	0.564	0.518	0.523	0.535
	(0.41)	(0.40)	(0.36)	(0.36)	(0.37)
Log (inside ownership)	0.523	0.511	0.503	0.560	0.545
	(1.42)	(1.40)	(1.38)	(1.53)	(1.48)
CEO chairman	-0.663	-0.672	-0.698	-0.660	-0.657
	(-0.88)	(-0.89)	(-0.93)	(-0.88)	(-0.88)
Acquirer's pre-announcement stock price run-up	-3.394***	-3.404***	-3.412***	-3.429***	-3.450***
	(-3.65)	(-3.70)	(-3.65)	(-3.67)	(-3.71)
Free cash flow	-5.492	-5.674	-5.850	-5.718	-5.613
	(-0.44)	(-0.47)	(-0.47)	(-0.46)	(-0.45)
Leverage	3.240	3.103	3.399	3.667	3.176
	(0.89)	(0.89)	(0.93)	(0.99)	(0.87)
Log (total asset)	0.285	0.273	0.193	0.236	0.224
	(1.13)	(1.12)	(0.78)	(0.94)	(0.88)
Log (firm age)	-0.185	-0.202	-0.246	-0.236	-0.241
	(-0.36)	(-0.40)	(-0.48)	(-0.47)	(-0.48)
Tobin's q	0.414**	0.416**	0.438**	0.398**	0.405**
	(2.30)	(2.33)	(2.45)	(2.20)	(2.24)
Intrastate	0.931	0.952	0.968	0.914	0.958
	(0.86)	(0.88)	(0.90)	(0.85)	(0.89)
Relative deal size	-3.416***	-3.420***	-3.303***	-3.369***	-3.358***
	(-3.08)	(-3.08)	(-3.04)	(-3.05)	(-3.07)
All-cash deal	1.849***	1.853***	1.844***	1.767***	1.771***
	(2.72)	(2.72)	(2.73)	(2.61)	(2.60)
Diversifying acquisitions	-0.785	-0.786	-0.745	-0.751	-0.773
	(-1.14)	(-1.14)	(-1.09)	(-1.10)	(-1.13)
Year dummies	Yes	Yes	Yes	Yes	Yes
R^2	0.14	0.14	0.15	0.15	0.14

Table 8: S&P 500-CEO Busy and Non-Busy Outside directors

Variable	(1)	(2)	(3)	(4)
Percentage of busy outside directors	-3.693**	-5.260***	-3.802**	-5.241***
	(-2.36)	(-3.00)	(-2.35)	(-2.87)
Percentage of SP500-CEO busy outside directors		8.142**		7.477**
		(2.43)		(2.21)
Percentage of SP500-CEO non-busy outside directors	3.053	3.756	5.060	5.705
	(0.57)	(0.71)	(0.92)	(1.04)
Percentage of outside directors	4.567*	4.136*	3.649	3.254
	(1.83)	(1.66)	(1.40)	(1.25)
Log (board size)	0.541	0.447	0.432	0.346
	(0.37)	(0.31)	(0.29)	(0.24)
Log (inside ownership)	0.529	0.567	0.561	0.597
	(1.43)	(1.55)	(1.49)	(1.59)
CEO chairman	-0.650	-0.644	-0.682	-0.676
	(-0.87)	(-0.86)	(-0.90)	(-0.89)
Acquirer's pre-announcement stock price run-up	-3.394***	-3.428***	-4.020***	-4.052***
	(-3.65)	(-3.67)	(-4.00)	(-4.01)
Free cash flow	-5.492	-5.721	-5.341	-5.551
	(-0.44)	(-0.46)	(-0.42)	(-0.43)
Leverage	3.228	3.658	2.563	2.958
	(0.88)	(0.99)	(0.68)	(0.78)
Log (total asset)	0.283	0.233	0.393	0.347
	(1.12)	(0.93)	(1.55)	(1.38)

Log (firm age)	-0.195	-0.249	-0.421	-0.470
	(-0.38)	(-0.49)	(-0.81)	(-0.91)
Tobin's q	0.409**	0.392**	0.423**	0.407**
	(2.29)	(2.19)	(2.24)	(2.15)
Intrastate	0.970	0.961	1.007	0.998
	(0.89)	(0.88)	(0.92)	(0.91)
Relative deal size	-3.456***	-3.418***	-3.793***	-3.758***
	(-3.15)	(-3.13)	(-3.34)	(-3.32)
All-cash deal	1.838***	1.751***	1.797***	1.717***
	(2.71)	(2.60)	(2.71)	(2.61)
Diversifying acquisitions	-0.792	-0.758	-0.640	-0.609
	(-1.15)	(-1.10)	(-0.93)	(-0.88)
Year dummies	Yes	Yes	Yes	Yes
R^2	0.14	0.15	0.17	0.17

Table 9: S&P 500- and Non-S&P 500-CEO Busy Outside Directors

Variable	(1)	(2)	(3)	(4)
Percentage of busy outside directors	-3.727**	-5.398***	-3.907**	-5.450***
	(-2.29)	(-2.90)	(-2.32)	(-2.81)
Percentage of SP500-CEO busy outside directors		8.203**		7.567**
		(2.39)		(2.17)
Percentage of non-SP500-CEO busy outside directors	0.147	1.882	1.226	2.827
	(0.03)	(0.35)	(0.23)	(0.51)
Percentage of outside directors	4.690*	4.222*	3.817	3.386
	(1.89)	(1.70)	(1.48)	(1.31)
Log (board size)	0.600	0.506	0.522	0.435
	(0.41)	(0.35)	(0.36)	(0.30)
Log (inside ownership)	0.523	0.563	0.553	0.590
	(1.41)	(1.53)	(1.45)	(1.56)
CEO chairman	-0.663	-0.662	-0.704	-0.703
	(-0.88)	(-0.89)	(-0.93)	(-0.93)
Acquirer's pre-announcement stock price run-up	-3.394***	-3.422***	-4.017***	-4.042***
	(-3.63)	(-3.64)	(-3.97)	(-3.98)
Free cash flow	-5.494	-5.743	-5.354	-5.584
	(-0.44)	(-0.46)	(-0.42)	(-0.43)
Leverage	3.246	3.759	2.637	3.110
	(0.88)	(1.00)	(0.69)	(0.80)
Log (total asset)	0.285	0.241	0.400	0.359
	(1.13)	(0.96)	(1.57)	(1.42)
Log (firm age)	-0.185	-0.235	-0.402	-0.448
	(-0.36)	(-0.46)	(-0.77)	(-0.87)
Tobin's q	0.414**	0.402**	0.434**	0.422**
	(2.29)	(2.21)	(2.26)	(2.19)
Intrastate	0.930	0.892	0.928	0.893
	(0.85)	(0.82)	(0.84)	(0.81)
Relative deal size	-3.416***	-3.374***	-3.730***	-3.691***
	(-3.08)	(-3.05)	(-3.23)	(-3.20)
All-cash deal	1.850***	1.775***	1.822***	1.754***
	(2.70)	(2.61)	(2.71)	(2.63)
Diversifying acquisitions	-0.784	-0.734	-0.620	-0.573
	(-1.14)	(-1.07)	(-0.90)	(-0.83)
Year dummies	Yes	Yes	Yes	Yes
R^2	0.14	0.15	0.17	0.17

On the Implications of the Increased Presence of Foreign Banks in Central America, Panama and the Dominican Republic

Abstract

Foreign banks have made a decision to increase their presence in Central America. This decision is a result of the changing political and economic environment in the area.. This strategic decision reflects a desire to take advantage of the banking opportunities available due to the region□?Ts integration through the Central American Free Trade Agreement and favorable economic growth in the region. This presence also reflects an improved business environment due to structural reforms and changing macroeconomic policies. The purpose of this paper is to address the different issues arising from this increased presence.

An Examination of India's Stock and Bond Markets Co-movements

Bharat Kolluri and Mahmoud Wahab
Respectively, Professors of Economics and Finance
Department of Economics, Finance and Insurance
Barney School of Business
University of Hartford, Connecticut, U.S.A.

Paper Presented at the Northeast Decision Sciences Institute
2010 Annual Conference, Washington D.C.
March 26-March 28, 2010

An Examination of India's Stock and Bond Markets Co-movements

Abstract

We examine the relationship between India's stock and bond markets using several multivariate specifications that vary by: (a) whether or not conditional heteroskedasticity is explicitly modeled in a multivariate GARCH framework, and (b) whether or not additional variables (beyond India's own stock and bond returns) such as foreign stock market variables and their corresponding conditional volatilities are included. We find that: (a) the U.K. stock market is the most influential market on India's capital markets, while the effect of the U.S. stock market is minimal, (b) conditional stock market volatility both in India and outside of India positively affects India's bond returns; higher Indian and foreign stock market volatilities lead to higher Indian bonds returns, (c) the effects of Japan and China on India's capital markets are limited. This preliminary evidence suggests that stock market volatility both inside and outside of India trigger higher returns on India's bond market; evidence that suggests a flight-to-quality phenomenon in effect as capital moves out of high risk assets to safe assets.

An Examination of India's Stock and Bond Markets Co-movements

I. Introduction

We examine linkages between India's Stock and Bond Markets during the most recent time period from 1997 to 2009. We estimate Vector Auto-Regression (VAR) models using weekly returns. The VAR models are estimated with and without accounting for conditional heteroskedasticity. Further, variations of the VAR model are estimated by augmenting the base specification with additional variables beyond lagged stock and bond returns in India such as: (a) returns on the three largest capitalization equity markets in the world: U.S., U.K., and Japan, (b) returns on China's stock market (China is the second largest trading partner with India after the U.S.), and (c) returns on an index of Emerging Asia stock markets to proxy for regional capital markets conditions, and (d) conditional volatilities of alternative stock return indices.

In comparison to other Asian capital markets, India's capital markets have received little attention. While there are scattered studies dealing with co-movement of India's stock market with other stock markets, there is virtually a complete lack of evidence on the interrelationship between India's stock and bond markets, an issue of importance to Indian and foreign investors alike seeking diversified portfolios. The purpose of this exploratory study is to investigate the relationship between India's stock and bond markets, particularly in terms of: (a) how each market affects the other, and (b) how changes in both markets are driven by movements in other (developed and emerging) markets.

One of the earlier studies on co-movements of India's stock market with other markets is due to Sharma and Kennedy (1977) who examined co-movements of India's stocks with London and New York markets. The authors found no evidence of any statistically significant co-movement patterns, and attributed this to the closed nature of the Indian economy during the period leading up to the 1970s and, particularly, presence of heavy controls on capital flows in and out of India, which did not get lifted till the 1980s (Wong, Agarwal and Du, 2005). Beginning with the establishment of the Securities and Exchange Board of India (SEBI) in 1988, and India's stock market reforms in 1991, India has increasingly opened up its economy and capital markets to foreign investors, and has encouraged both foreign direct and indirect (securities) investments. Indeed, India is evolving as one of the world's largest recipients of foreign capital from major developed countries of the U.S., U.K., and Japan (Wong et al., 2005). Furthermore, in recent years, India has witnessed a remarkable surge in trade with China so that by the end of 2004, India's total trade with China neared $14 billion, second only to the U.S. (Chen et al., 2005). Despite the opening up of India's economy and capital markets and its growing visibility in world trade and finance, the volume of research on India's capital markets remains meager. This lack of attention to India's capital markets surprisingly comes at a time when the Bombay Stock Exchange (BSE) has become the world's largest stock Exchange in terms of number of listed companies, and is ranked among the top ten markets in the world in terms of market capitalization. As of end of 2007, the BSE's market capitalization reached 1.8 Trillion U.S. dollars (Raj and Dahl, 2008).

We jointly estimate conditional mean return models for India's stock and bond markets as a function of: (a) their own lags, (b) returns on major foreign stock markets, and (c) conditional volatilities of India's stock returns and volatilities of foreign stock returns. The models are estimated under both a conditional homoskedasticity assumption, as well as with explicit modeling of time-varying conditional volatility using a *VAR-MGARCH* (Vector Auto-regression Multivariate Generalized autoregressive conditional heteroskedasticity) model. A *VAR-MGARCH* model jointly estimates conditional expected returns and variances to enhance efficiency of the parameter estimates. To parameterize conditional covariances, we adopt a Constant Conditional Correlation (CCC) specification (Bollerslev, 1990) of conditional heteroskedasticity in which time-varying volatility of each market is described in terms of only its own lagged squared residuals and lagged volatility. A constant correlation specification is deemed adequate for our purpose as our primary interest lies in co-movements of expected returns on India's stock and bond markets more so than in co-movements of their corresponding volatilities. To capture the effect of conditional stock return volatility on stocks and bonds, however, we use estimated time-varying Indian stock market volatility and foreign stock markets volatilities in India's stock and bond return equations. Including these volatility variables may help us understand the role that local stock market or foreign stock market volatilities have on asset substitution between stocks and bonds in India. To the authors' knowledge, this is the first study that addresses the joint relationship between India's stock and bond markets returns.

Chen, Lobo, and Wong (2005) used weekly data for the period from 1991 and 2004 to examine three bilateral stock market relationships involving India, U.S. and China. Their evidence can be summarized as follows: (a) the Indian stock market adjusts to departures from long-term equilibrium with both U.S. and China, while the Chinese market adjusts to disequilibrium with only the U.S. market, (b) the U.S. stock market leads the Indian market in volatility (but not returns), but leads the Chinese market in returns, but not volatility (c) the Indian and Chinese markets are interactive in both returns and volatility; while China leads India in returns, India leads China in volatility, and both markets seem to be more closely linked with each other than they are with the U.S.

In a different study, Wong, Agarwal and Du (2005), used weekly data for the period between 1991 and 2003 to examine relationships between India's stock market and stock markets of the U.S., U.K., and Japan. They reported that: (a) the India's stock market was cointegrated with all three developed stock markets, (b) there is a unidirectional causality from U.S. and Japanese stock markets (but not from the U.K. market) to India's stock market, and (c) no reverse causality is detected from India's stock market to any of the three developed stock markets. Raj and Dhal (2008) examined co-movements between India's stock market and developed stock markets of the U.S., U.K., and Japan in addition to two major regional stock markets of Hong Kong and Singapore.. They used daily and weekly returns in both U.S. dollar-adjusted returns as well as (unadjusted) Indian Rupee returns for the period between 1993 and 2008. They found that while U.K. returns have a positive spillover effect on India's stocks, the impact of U.S. stock returns is negative. The impact of developed stock markets on India's stock market was found to be more pronounced with weekly than with daily data, while the impact of established regional stock markets of Hong Kong and Singapore on India was more pronounced on a daily than weekly basis. Japan's stock market had little to no effect on India's stock market compared to U.S. and U.K. stock markets.

As is well known, uncertainty in equity markets is often reflected in government bond returns through the so-called "flight to quality" phenomenon, whereby investors move capital from stocks into risk-free bonds during turbulent times (Stivers and Sun, 2002). For example, in the case of the U.S., while the (unconditional) correlation between U.S. stocks and bonds is, on average, positive (roughly at 0.3), it has been shown to decline and even turn negative during uncertain times in the U.S. stock market as a result of substitution of risk-free assets (government bonds) for risky assets (stocks). This results in an outflow of funds from stocks into bonds causing price pressures in opposite directions in both markets. Stivers and Sun (2002) found that stock market uncertainty plays an important role in explaining changing U.S. stock-bond correlations over time.

Because India along with China are both included in the Emerging Asian markets stock index used in the study to proxy for smaller regional Asia markets conditions, we orthogonalize India's and China's stock market returns that are part of the index by regressing that index's returns onto both the Bombay Stock Exchange (BSE) returns and the Shanghai Stock Index returns (China's Index) then use the residuals from this regression as a proxy for Emerging Asia stock markets index that is devoid of the effects of both India and China. This orthogonalization procedure accomplishes two objectives in one step: first it excludes India's stock market component from the index in order to focus on only (regional Asia) shocks that are external to India's capital markets, and second, by excluding China from the Emerging Asia index we isolate China's effects on India's capital markets from the effects of the rest of Emerging Asia stock markets given the special trading partner status that China enjoys with India (China is India's second largest trading partner).

We estimate various multivariate models using: (a) Seemingly-Unrelated Regression (SUR), and (b) and multivariate *VAR(p)-MGARCH(p,q)-CCC* model that explicitly models time-varying volatility in India's stock and bond markets. Because a constant conditional correlation specification assumes that volatility of a market is described only by its own lagged squared residuals and volatilities, it does not permit an investigation of volatility spillovers between the two markets. However, this should not be of major concern as: (a) it is fairly well known that volatility spillovers occur from stocks to bonds (not the other way around), and (b) our focus in this paper is on interactions of stock and bond returns in India, and not on interactions of their volatilities.

Two conditional mean equations (one for India's stock returns, the other for India's bond returns) are estimated. Similarly, two conditional variance equations in a *VAR(p)-MGARCH(p,q)* model framework are estimated, one for India's stocks, while the other is for its bonds. Alternative versions of the conditional mean return equations are estimated to examine robustness of the results. The basic equation is a VAR(p) that includes only India's stock and bond returns. We gradually augment these equations with: (a) returns on major foreign equity markets of the U.S., U.K., Japan, and China, as well as an index of Emerging stock markets, and (b) expected stock market volatility variables estimated from univariate GARCH(p,q) models for each of the stock market indices (India's stock index included). Stock volatility variables are included to examine the effects that local, regional and global stock returns uncertainty have on India's stock and bond returns. Previous studies on India's stock market have included only foreign stock returns. This study is the first to link India's stock and bond market returns to explore their joint relationship and their relationship with returns on other stock markets and their corresponding volatilities. The role of stock market volatility in triggering asset re-allocation between India's stock and bond markets is examined. Since these stock indices are not expected to be highly correlated, we are not too concerned about collinearity issues in the conditional mean returns equations. To summarize, the primary issues explored in this paper include: (a) investigating the extent to which movements in India's stock and bond returns are predictable by each other's movements as well as movements in regional and global stock markets, and (b) investigating the extent to which local, regional and global stock market uncertainty explain movements in India's stock and bond returns.

II. Data and Models

The stock indices used in this study are the Bombay Stock Exchange Index (BSE), the Standard and Poor's S&P500 Index of the U.S., the Tokyo Stock Exchange Index (TOPIX) of Japan, the Shanghai Stock Exchange Index of China, the Financial Times Stock Exchange Index of the U.K., and an Emerging Asian stock markets index to represent a regional stock markets factor. All data are from Global Financial Data (www.globalfinancialdata.com). Weekly Wednesday-to-Wednesday closing prices are used to avoid a variety of well known biases such as day-of-the-week effects (volatilities tend to be higher on Fridays and Mondays, and returns tend to be negative on Mondays, compared to other days of the week). In addition, use of weekly data avoids problems relating to stale prices arising from non-synchronous trading and time-zone differences of markets that do not fully or partially overlap in trading hours. As is well known, daily returns tend to be too noisy, while monthly returns tend to be too coarse; accordingly, weekly data strike a good balance between the two data frequencies. All models are estimated using U.S. dollar-adjusted returns on all variables to make results relevant to a U.S. investor. Nominal dollar-adjusted returns are computed as: $\{R_{i,N,\$,t} = (1+R_{i,N,f,t})*(1+\%\Delta S_N(\$,fc))-1\}$, where $R_{i,N,\$,t}$. are nominal dollar returns on market (i), $R_{i,N,f,t}$ are nominal returns in local currency, and $\%\Delta S_N(\$,fc)$ are changes in nominal spot exchange rates using New York market mid-point {(bid+ask)/2} quotes for each foreign currency on every Wednesday.

First, we estimate a multivariate Seemingly Unrelated (*SUR-VAR(p)*) model of India's stock and bond returns without explicitly modeling conditional heteroskedasticity, as follows:

$$R_{s,t} = \alpha_{s,0} + \sum_{j=1}^{3} \beta_{ss,j} R_{s,t-j} + \sum_{j=1}^{3} \beta_{sb,j} R_{b,t-j} + \varepsilon_{s,t} \tag{1.a}$$

$$R_{b,t} = \alpha_{b,0} + \sum_{j=1}^{3} \beta_{bb,j} R_{b,t-j} + \sum_{j=1}^{3} \beta_{bs,j} R_{s,t-j} + \varepsilon_{b,t} \tag{1.b}$$

where $R_{s,t}$ and $R_{b,t}$ denote, respectively, stock and bond returns. Eqs. (*1.a*) and (*1.b*) represent the base specification. The optimal lag order is found to be 3 based on Akaike's Information Criterion corrected for degrees of freedom (AICC). We augment eqs. (*1.a*) and (*1.b*) by including foreign stock market variables together with India's 90 government T.bill rates (used to proxy for local business conditions in India), still without explicitly modeling conditional heteroskedasticity, as follows:

$$\begin{aligned} R_{s,t} = {} & \alpha_{s,0} + \alpha_{s,1} R_{s,t-1} + \beta_{0,b} R_{b,t-1} + \sum_{j=0}^{1} \beta_{1,s,j} R_{us,t-j} + \sum_{j=0}^{1} \beta_{2,s,j} R_{ja,t-j} + \\ & \sum_{j=0}^{1} \beta_{3,s,j} R_{uk,t-j} + \sum_{j=0}^{1} \beta_{4,s,j} R_{ch,t-j} + \sum_{j=0}^{1} \beta_{5,s,j} R_{EM,t-j} + \sum_{j=0}^{1} \beta_{6,s,j} TBY_{t-j} + \varepsilon_{s,t} \end{aligned} \tag{2.a}$$

$$\begin{aligned} R_{b,t} = {} & \alpha_{b,0} + \alpha_{b,1} R_{b,t-1} + \beta_{0,s} R_{s,t-1} + \sum_{j=0}^{1} \beta_{1,b,j} R_{us,t-j} + \sum_{j=0}^{1} \beta_{2,b,j} R_{ja,t-j} + \\ & \sum_{j=0}^{1} \beta_{3,b,j} R_{uk,t-j} + \sum_{j=0}^{1} \beta_{4,b,j} R_{ch,t-j} + \sum_{j=0}^{1} \beta_{5,b,j} R_{EM,t-j} \sum_{j=0}^{1} \beta_{6,b,j} TBY_{t-j} + \varepsilon_{b,t} \end{aligned} \tag{2.b}$$

where $(\varepsilon_{s,t}, \varepsilon_{b,t})'$ is the residuals vector which is assumed to be conditionally normal and homoskedastic: $\boldsymbol{\varepsilon_{i,t} \mid \Omega_{i,t-1} \sim N(0,H)}$, and $R_{us,t}, R_{ja,t}, R_{uk,t}, R_{ch,t}, R_{EM,t}$ denote, respectively, U.S., Japan, U.K., Japanese, China and Emerging markets stock index returns. Standard errors for the SUR model are estimated using White's (1980) heteroskedasticity-consistent residuals' variance-covariance matrix, and TBY is the 90-day Indian government Treasury bill rate. Three more sets of models are estimated while explicitly modeling time-varying residuals' variances. The first is given by the system of eqs. (3.a)-(3.c), as follows:

$$R_{s,t} = \alpha_{s,0} + \sum_{j=1}^{3} \beta_{ss,j} R_{s,t-j} + \sum_{j=1}^{3} \beta_{sb,j} R_{b,t-j} + \varepsilon_{s,t} \tag{3.a}$$

$$R_{b,t} = \alpha_{b,0} + \sum_{j=1}^{3} \beta_{bb,j} R_{b,t-j} + \sum_{j=1}^{3} \beta_{bs,j} R_{s,t-j} + \varepsilon_{b,t} \tag{3.b}$$

$$\sigma^2_{s,t} = \sigma^2_s + b_{ss}\sigma^2_{s,t-1} + c_{ss}\varepsilon^2_{s,t-1} \text{ and } \sigma^2_{b,t} = \sigma^2_b + b_{bb}\sigma^2_{b,t-1} + c_{bb}\varepsilon^2_{b,t-1};$$

$$\sigma_{s,b,t} = \sigma_{b,s,t} = \rho_{s,b}(\sigma^2_{s,t} * \sigma^2_{b,t})^{1/2}, \text{ so that } \rho_{s,b} = \sigma_{s,b,t} / \sqrt{\sigma^2_{s,t} . \sigma^2_{b,t}} \qquad (3.c)$$

A lag order of 3 on the right-hand side of eqs. (1.a) and (1.b) was found to be optimal based on the minimum AICC. The second model augments eqs. (*3.a*)-(*3.c*) by adding foreign stock market variables of the U.S., Japan, U.K., China and an Emerging stock markets index in addition to the Indian government's 90-day T.bill rate. The optimal lag order based on the minimum AICC criterion was found to be one. Therefore, we estimate the following second set of models given in eqs. (*4.a*)-(*4.c*):

$$R_{s,t} = \alpha_{s,0} + \alpha_{s,1}R_{s,t-1} + \beta_{0,b}R_{b,t-1} + \sum_{j=0}^{1}\beta_{1,s,j}R_{us,t-j} + \sum_{j=0}^{1}\beta_{2,s,j}R_{ja,t-j} +$$
$$\sum_{j=0}^{1}\beta_{3,s,j}R_{uk,t-j} + \sum_{j=0}^{1}\beta_{4,s,j}R_{ch,t-j} + \sum_{j=0}^{1}\beta_{5,s,j}R_{EM,t-j} + \sum_{j=0}^{1}\beta_{6,s,j}TBY_{t-j} + \varepsilon_{s,t} \qquad (4.a)$$

$$R_{b,t} = \alpha_{b,0} + \alpha_{b,1}R_{b,t-1} + \beta_{0,s}R_{s,t-1} + \sum_{j=0}^{1}\beta_{1,b,j}R_{us,t-j} + \sum_{j=0}^{1}\beta_{2,b,j}R_{ja,t-j} +$$
$$\sum_{j=0}^{1}\beta_{3,b,j}R_{uk,t-j} + \sum_{j=0}^{1}\beta_{4,b,j}R_{ch,t-j} + \sum_{j=0}^{1}\beta_{5,b,j}R_{EM,t-j} + \sum_{j=0}^{1}\beta_{6,b,j}TBY_{t-j} + \varepsilon_{s,t} \qquad (4.b)$$

$$\sigma^2_{s,t} = \sigma^2_s + b_{ss}\sigma^2_{s,t-1} + c_{ss}\varepsilon^2_{s,t-1} \text{ and } \sigma^2_{b,t} = \sigma^2_b + b_{bb}\sigma^2_{b,t-1} + c_{bb}\varepsilon^2_{b,t-1};$$

$$\sigma_{s,b,t} = \sigma_{b,s,t} = \rho_{s,b}(\sigma^2_{s,t} * \sigma^2_{b,t})^{1/2}, \text{ so that } \rho_{s,b} = \sigma_{s,b,t} / \sqrt{\sigma^2_{s,t} . \sigma^2_{b,t}} \qquad (4.c)$$

Lastly, the third set of models augments eqs. (4.a)-(4.c) with expected stock volatility variables for India, U.S., Japan, U.K., China, and Emerging Stock Markets Index. Expected time-varying volatility is estimated as predicted conditional variances from a GARCH(*p,q*) model as follows: $\sigma^2_{i,t} = \sigma^2_i + b_{ii}\sigma^2_{i,t-1} + c_{ii}\varepsilon^2_{i,t-1}$; $\forall$ *i=(U.S., Japan, U.K., China, and an Emerging Asia Stock Markets Indices).* Estimation is conducted using maximum likelihood. Eqs. (*5.a*)-(*5.c*) are given as follows:

$$R_{s,t} = \alpha_{s,0} + \alpha_{s,1}R_{s,t-1} + \beta_{0,b}R_{b,t-1} + \sum_{j=0}^{1}\beta_{1,s,j}R_{us,t-j} + \sum_{j=0}^{1}\beta_{2,s,j}R_{ja,t-j} + \sum_{j=0}^{1}\beta_{3,s,j}R_{uk,t-j} +$$
$$\sum_{j=0}^{1}\beta_{4,s,j}R_{ch,t-j} + \sum_{j=0}^{1}\beta_{5,s,j}R_{EM,t-j} + \sum_{j=0}^{1}\beta_{6,s,j}TBY_{t-j} + \sum_{j=0}^{1}\gamma_{s,ind,t-j}\sigma^2_{ind,t-j} + \sum_{j=0}^{1}\gamma_{s,us,j}\sigma^2_{us,t-j} + \qquad (5.a)$$
$$\sum_{j=0}^{1}\gamma_{s,ja,j}\sigma^2_{ja,t-j} + \sum_{j=0}^{1}\gamma_{s,uk,j}\sigma^2_{uk,t-j} + \sum_{j=0}^{1}\gamma_{s,ch,j}\sigma^2_{ch,t-j} + \sum_{j=0}^{1}\gamma_{s,EM,j}\sigma^2_{EM,t-j}$$

$$R_{b,t} = \alpha_{b,0} + \alpha_{b,1}R_{b,t-1} + \beta_{0.s}R_{s,t-1} + \sum_{j=0}^{1} \beta_{1,b,j}R_{us,t-j} + \sum_{j=0}^{1} \beta_{2,b,j}R_{ja,t-j} +$$
$$\sum_{j=0}^{1} \beta_{3,b,j}R_{uk,t-j} + \sum_{j=0}^{1} \beta_{4,b,j}R_{ch,t-j} + \sum_{j=0}^{1} \beta_{5,b,j}R_{EM,t-j} + \sum_{j=0}^{1} \beta_{6,b,j}TBY_{t-j} +$$
$$\sum_{j=0}^{1} \gamma_{b,ind,t-j}\sigma^2{}_{ind,t-j} + \sum_{j=0}^{1} \gamma_{b,us,j}\sigma^2{}_{us,t-j} + \sum_{j=0}^{1} \gamma_{b,ja,j}\sigma^2{}_{ja,t-j} + \sum_{j=0}^{1} \gamma_{b,uk,j}\sigma^2{}_{uk,t-j} +$$
$$\sum_{j=0}^{1} \gamma_{b,ch,j}\sigma^2{}_{ch,t-j} + \sum_{j=0}^{1} \gamma_{b,EM,j}\sigma^2{}_{EM,t-j} + \varepsilon_{b,t} \qquad (5.b)$$

$(\varepsilon_{s,t}, \varepsilon_{b,t})' | \Omega_{t-1} \sim N(0,H_t)$ and

$\sigma^2{}_{s,t} = \sigma^2{}_s + b_{ss}\sigma^2{}_{s,t-1} + c_{ss}\varepsilon^2{}_{s,t-1}$ and $\sigma^2{}_{b,t} = \sigma^2{}_b + b_{bb}\sigma^2{}_{b,t-1} + c_{bb}\varepsilon^2{}_{b,t-1}$

$$\sigma_{s,b,t} = \sigma_{b,s,t} = \rho_{s,b}(\sigma^2{}_{s,t} * \sigma^2{}_{b,t})^{1/2} \quad \text{so that} \quad \rho_{s,b} = \sigma_{s,b,t} / \sqrt{\sigma^2{}_{s,t} . \sigma^2{}_{b,t}} \qquad (5.c)$$

A lag order of one proved to be optimal based on AICC. All returns are log price- relatives.

III. Empirical Results

Panel (A) of Table (1) presents estimates of the base model given in eqs. (*1.a*) and (*1.b*) with significance levels estimated using White's (1980) heteroskedasticity-correction. As seen, the only statistically significant variable in the stock returns equation is the frst lagged bond returns variable with a coefficient of 0.2359 (statistically significant at the 5 percent level). The bond returns equations shows a statistically significant own-lagged adjustment (coefficient estimate is -0.1728 which is statistically significant at the 1 percent level). There is evidence of a lagged adjustment at the third lag in India's bond returns to stock returns in India (the coefficient estimate is 0.0315, negligible, but statistically significant at the 10 percent level). Lagged adjustments in stock and bond markets may not necessarily be indicative of pricing inefficiencies if expected returns in both markets are time-varying. Alternatively, they may be indicative of a misspecification error due to omitted variables. The R^2 coefficient of the model is 0.241. To summarize, India's stock and bond returns exhibit some feedback on a lagged basis.

Panel (B) in Table (1) presents estimates from an augmented model which includes stock market returns from foreign markets. The only statistically significant variable in the stock returns equation is own lagged returns (coefficient=0.0741 which is statistically significant at the 10 percent level). The bond returns equation shows evidence of more statistically significant relationships. For example, own-lagged bond returns variable has a coefficient of -0.1879 (statistically significant at the 1 percent level). Out of all foreign stock markets, the U.K. stock markets exerts the strongest influence on a weekly returns basis (contemporaneous coefficient estimate=0.1838 which is statistically significant at the 1 percent level) followed by a one period-lagged effect from the U.S. stock market (coefficient estimate=-0.0866 which is statistically significant at the 10 percent level). The coefficients on contemporaneous China and Emerging Asia indices, although statistically significant, are negligible in magnitude. In summary, there is some evidence of own lagged partial adjustments in India's stock and bond markets. Further, there is evidence that bond returns Granger-cause stock returns at the first lag, while stock returns Granger-cause bond returns at the third week (the former's magnitude is double the latter's magnitude). Lagged adjustments are not necessarily indicative of market inefficiencies as they may equally imply time-varying risk premiums, or worse, some form of model misspecification. The U.K. stock market exerts a positive influence on India's bond market, while the U.S. stock market exerts a negative influence; however, the impact of the U.K. is twice that of the U.S.

Estimates of eqs. (3.a)-(3.c) are reported in Panel (A) of Table (2), while estimates of eqs. (4.a)-(4.c) are reported in Panel (B). Both sets of models explicitly accommodate conditional heteroskedasticity. Eqs. (3.a)-(3.c) do not include foreign stock market returns or volatility variables. As seen, the bond equation shows more statistically significant coefficients than the stock equation. In particular, there is evidence of: (a) lagged partial adjustments in India's bond market that stretches out to all three lags (being statistically significant at the 1 percent level for the first two lags, and at the 10 percent level for the third lag), (b) stock returns lead bond returns up to two lags with negative coefficients suggesting that when stock returns are negative, India's bond returns are positive (evidence that corroborates an asset re-allocation story between India's stock and bond markets). The ARCH and GARCH coefficients in the multivariate volatility portion of the model are statistically significant and their sums are less than unity which suggests stationary volatility for India's stock and bond returns. Panel (B) presents estimates of augmented model of eqs. (4.a)-(4.c) that include foreign stock returns variables. Once more, the only foreign market that seems to affect India's stock market is the U.K. market (coefficient estimate=0.1581 which is statistically significant at the 5 percent level). China's stock market exerts very minimal influence on India's bond market, and does not correlate with India's stock market. The coefficients on India's contemporaneous and one-week lagged 90-day T.bill yields are negative and statistically significant at the 1 percent level (as expected, since higher short term interest dampen economic activity and the stock market). ARCH and GARCH effects are statistically significant only for the stock returns equation, and only ARCH effects are statistically significant for the bond returns equation (indicating short memory in the conditional volatility process of bond returns that is contrary to the long memory process of conditional stock returns volatility in the Indian stock market). The estimated conditional correlation coefficient is 0.1517 which is less than the unconditional correlation reported earlier in Table (2) of around 0.24. Therefore, it appears that the VAR(1)-MGARCH(1,1) model has successfully explained some of the reported unconditional cross-correlation.

Lastly, we report in Table (3) estimates of the broadest specification that includes: own returns, foreign stock returns, and conditional volatilities of stock returns from all six stock markets. Japan's stock market movements correlate contemporaneously with India's stock movements, but the stronger co-movement is still with the U.K. stock market, both contemporaneously and with a one week-lag. A part of India's stock market responses to movements in the U.K. stock market occur the same week, with additional adjustments taking place the following week. Indian government Treasury bill yields have no effect on India's stock returns. India's stock market volatility affects its returns on a contemporaneous basis suggesting that higher stock market volatility leads to higher required returns. Volatilities of Japanese equities, U.K. equities and Emerging markets' equities also affect stock returns in India. India's bond market is moved by its stock market as well as stock markets of the U.K. and Japan. What is particularly interesting is evidence of the effect of stock market volatility on India's bond market. More specifically, there is evidence that last week's stock market volatility in India leads to higher bond market returns the following week, which may indicate substitution of bonds for stocks whenever stock volatility picks up. By the same token, volatility in Emerging Asia stock markets also leads to higher Indian bond market returns through a flight to safety in the Indian government bonds. ARCH and GARCH effects in India's stock market are statistically significant, but, once more, only ARCH effects are statistically significant for India's bond market. The estimated conditional correlation between India's stock and bond markets is 0.1928 (and is statistically significant at the 1 percent level). Diagnostics tests (not reported, but available upon request) suggest that the specification given in eqs. (5.a)-(5.c) capture fairly well the dynamics of India's stock and bond markets returns and volatilities.

IV. Summary and Conclusions

We examine the relationship between India's stock market and bond market returns using the 10-year Indian Government bond. Results suggest that the most influential foreign stock market for India's capital markets is the U.K. with the U.S. market exerting little to no influence. Subsequently, we model conditional heteroskedasticity by estimating a number of

VAR(p)-MGARCH(p,q) models that differ in the way we specify conditional expected returns. We start with a bivariate model of only India's stock and bond markets, followed by models that include foreign stock market variables, and end with the broadest specification that includes foreign stock returns and their corresponding conditional volatilities. We find that: (a) the most influential market on India's capital markets is still the U.K., (b) contemporaneous and lagged conditional stock market volatilities in India, U.K., and Emerging Asia markets index are positively correlated with returns on Indian bonds, evidence we interpret as indicative of substitution of risky stocks for risk-free bonds. This is evidence which is similar to U.S. evidence that U.S. stock market volatility drives up bond returns through a flight-to-quality and/or cross-hedging trades of stock market risk. The study can be extended in several directions. For example, it would be of interest to examine the effects of alternative conditional heteroskedasticiy specifications on the reported relationships between stock volatility and bond returns in India. It would also be of interest to re-examine these relationships in a trivariate framework that jointly models India's stock market, bond market, and money markets to determine if a substitution effect from risky to risk-free assets goes not only into India's government bond market, but also into India's very short term money market as well. Last, but not least, it would be of interest to estimate these relationships in Indian Rupee equivalent returns to examine stability of the results regardless of choice of currency. It is fairly well established that the covariance between foreign exchange rates changes and stock returns is empirically small, while the covariance between exchange rate changes and bond returns is much larger. Whether or not the estimated relationships hold in rupee equivalents is an empirical question.

Table (1)
Parameter Estimates of
VAR (p) Model With White's Heteroskedasticity-Correction
(Jan. 1997-Dec. 2009)

Panel A: Estimates of Eqs. (1.a)-(1.b)

$\alpha_{S,0}$ -0.0086 (0.279)	$\beta_{SS,1}$ 0.0559 (0.195)	$\beta_{SS,2}$ 0.0380 (0.380)	$\beta_{SS,3}$ 0.0194 (0.655)	$\beta_{Sb,1}$ 0.2359 (0.021)**	$\beta_{Sb,2}$ -0.0170 (0.866)
$\beta_{Sb,3}$ 0.0171 (0.867)	$\alpha_{b,0}$ 0.0021 (0.533)	$\beta_{bb,1}$ -0.1728 (<0.0001)***	$\beta_{bb,2}$ 0.0365 (0.399)	$\beta_{bb,3}$ -0.0624 (0.144)	$\beta_{bs,1}$ 0.0121 (0.506)
$\beta_{bs,2}$ -0.0043 (0.815)	$\beta_{bs,3}$ 0.0315 (0.085)*	System R^2=0.241			

Panel B: Estimates of Eqs. (2.a)-(2.b)

$\alpha_{S,0}$ 0.0017 (0.916)	$\alpha_{S,1}$ 0.0741 (0.085)*	$\beta_{0,b}$ 0.2773 (0.008)***	$\beta_{1,s,0}$ 0.0816 (0.464)	$\beta_{1,s,1}$ -0.0811 (0.475)	$\beta_{2,s,0}$ 0.0634 (0.414)
$\beta_{2,s,1}$ -0.1197 (0.111)	$\beta_{3,s,0}$ 0.0655 (0.317)	$\beta_{3,s,1}$ -0.0434 (0.504)	$\beta_{4,s,0}$ -0.0007 (0.810)	$\beta_{4,s,1}$ -0.0019 (0.497)	$\beta_{5,s,0}$ 0.0143 (0.786)
$\beta_{5,s,1}$ 0.0001 (0.849)	$\beta_{6,s,0}$ -5.4162 (0.762)	$\beta_{6,s,1}$ 1.7065 (0.517)	$\alpha_{b,0}$ -0.0008 (0.190)	$\alpha_{b,1}$ -0.1879 (<0.0001)***	$\beta_{0,s}$ 0.0151 (0.390)
$\beta_{1,b,o}$ -0.0398 (0.382)	$\beta_{1,b,1}$ -0.0866 (0.061)*	$\beta_{2,b,0}$ 0.1838 (<0.0001)***	$\beta_{2,b,1}$ -0.0831 (0.215)	$\beta_{3,b,0}$ 0.0119 (0.656)	$\beta_{3,b,1}$ -0.0311 (0.247)
$\beta_{4,b,0}$ 0.0003 (0.016)**	$\beta_{4,b,1}$ 0.000 (0.956)	$\beta_{5,b,0}$ 0.0053 (0.807)	$\beta_{5,b,1}$ 0.0004 (0.018)**	$\beta_{6,b,0}$ -1.7178 (0.814)	$\beta_{6,b,1}$ -0.1991 (0.978)
System R^2 = 0.249					

Notes: ****, ** and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively.*

Table (2)
Parameter Estimates of
VAR (p) – MGARCH (p, q) - CCC
(Jan. 1997-Dec. 2009)

Panel A: Estimates of Eqs. (3.a)-(3.c)					
$\alpha_{S,0}$ -0.0061 (0.0009)***	$\beta_{SS,1}$ 0.0640 (0.179)	$\beta_{SS,2}$ 0.0312 (0.512)	$\beta_{SS,3}$ 0.0251 (0.577)	$\beta_{Sb,1}$ 0.1175 (0.253)	$\beta_{Sb,2}$ 0.1529 (0.122)
$\beta_{Sb,3}$ -0.0108 (0.908)	$\alpha_{b,0}$ -0.0323 (0.0001)***	$\beta_{bb,1}$ 0.2401 (0.0001)***	$\beta_{bb,2}$ 0.2884 (0.0001)***	$\beta_{bb,3}$ -0.0684 (0.084)*	$\beta_{bs,1}$ -0.0321 (0.0055)***
$\beta_{bs,2}$ -0.0235 (0.025)**	$\beta_{bs,3}$ -0.007 (0.527)	σ_s^2 0.0002 (0.003)***	b_{ss} 0.7431 (0.0001)***	c_{ss} 0.1664 (0.0001)***	σ_b^2 0.0001 (0.0001)***
b_{bb} 0.0997 (0.0001)***	c_{bb} 0.3161 (0.0001)***	$\rho_{s,b}$ 0.1963 (0.0001)***			
Panel B: Estimates of Eqs. (4.a)-(4.c)					
$\alpha_{S,0}$ -0.0067 (0.006)***	$\alpha_{S,1}$ 0.0958 (0.045)**	$\beta_{0,b}$ 0.1435 (0.157)	$\beta_{1,s,0}$ 0.1010 (0.339)	$\beta_{1,s,1}$ -0.0706 (0.520)	$\beta_{2,s,0}$ 0.0905 (0.136)
$\beta_{2,s,1}$ -0.0704 (0.226)	$\beta_{3,s,0}$ 0.1581 (0.026)**	$\beta_{3,s,1}$ -0.0904 (0.208)	$\beta_{4,s,0}$ 0.0125 (0.481)	$\beta_{4,s,1}$ -0.0131 (0.459)	$\beta_{5,s,0}$ 0.0244 (0.635)
$\beta_{5,s,1}$ -0.0117 (0.826)	$\beta_{6,s,0}$ -1.0431 (0.468)	$\beta_{6,s,1}$ 0.8550 (0.491)	$\alpha_{b,0}$ -0.0011 (0.001)***	$\alpha_{b,1}$ 0.0323 (0.003)***	$\beta_{0,s}$ -0.0883 (0.071)*
$\beta_{1,b,o}$ -0.0406 (0.231)	$\beta_{1,b,1}$ 0.0071 (0.835)	$\beta_{2,b,0}$ 0.0202 (0.245)	$\beta_{2,b,1}$ -0.0281 (0.101)	$\beta_{3,b,0}$ 0.1181 (0.0001)***	$\beta_{3,b,1}$ -0.0396 (0.075)*
$\beta_{4,b,0}$ 0.0136 (0.017)**	$\beta_{4,b,1}$ -0.0046 (0.372)	$\beta_{5,b,0}$ 0.0100 (0.604)	$\beta_{5,b,1}$ 0.0281 (0.240)	$\beta_{6,b,0}$ -0.7611 (0.0001)***	$\beta_{6,b,1}$ -0.5969 (0.0001)***
σ_s^2 0.0014 (0.0101)*	b_{ss} 0.7476 (0.0001)***	c_{ss} 0.1549 (0.0007)***	σ_b^2 0.0001 (1.000)	b_{bb} 0.0003 (0.998)	c_{bb} 0.9035 (0.0001)***
$\rho_{s,b}$ 0.1571 (0.0003)***					
Notes: ****, ** and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively.*					

Table (3)
Parameter Estimates
Of VAR (p)-MGARCH (p, q)-CCC
Specification of Eqs. (5.a)-(5.c)

$\alpha_{s,0}$ -0.0071 (0.011)**	$\alpha_{s,1}$ 0.0367 (0.527)	$\beta_{0,b}$ 0.1260 (0.303)	$\beta_{1,s,0}$ 0.1027 (0.380)	$\beta_{1,s,1}$ -0.0950 (0.435)	$\beta_{2,s,0}$ 0.119 (0.074)*
$\beta_{2,s,1}$ -0.0432 (0.477)	$\beta_{3,s,0}$ 0.2256 (0.004)***	$\beta_{3,s,1}$ -0.1744 (0.027)**	$\beta_{4,s,0}$ 0.0039 (0.843)	$\beta_{4,s,1}$ -0.0031 (0.888)	$\beta_{5,s,0}$ 0.0039 (0.938)
$\beta_{5,s,1}$ 0.0042 (0.936)	$\beta_{6,s,0}$ -1.2025 (0.542)	$B_{6,s,1}$ 1.0328 (0.381)	$\gamma_{s,ind,0}$ 2.4756 (0.014)**	$\gamma_{s,ind,1}$ -2.1818 (0.055)*	$\gamma_{s,us,0}$ 1.9024 (0.472)
$\gamma_{s,us,1}$ 1.2671 (0.518)	$\gamma_{s,ja,0}$ 2.4316 (0.719)	$\gamma_{s,ja,1}$ 2.3150 (0.002)***	$\gamma_{s,uk,0}$ 2.8252 (0.035)**	$\gamma_{s,uk,1}$ 2.2501 (0.719)	$\gamma_{s,ch,0}$ 0.0002 (0.990)
$\gamma_{s,ch,1}$ 0.0008 (0.988)	$\gamma_{s,em,0}$ 3.5708 (0.0043)***	$\gamma_{s,em,1}$ 2.4418 (0.0013)***	$\alpha_{b,0}$ -0.0077 (0.0001)***	$\alpha_{b,1}$ -0.1110 (0.134)	$\beta_{0,s}$ (0.0169) (0.143)
$\beta_{1,b,0}$ -0.0327 (0.389)	$\beta_{1,b,1}$ 0.0122 (0.723)	$\beta_{2,b,0}$ 0.0225 (0.200)	$\beta_{2,b,1}$ -0.0350 (0.050)	$\beta_{3,b,0}$ 0.1273 (0.0001)	$\beta_{3,b,1}$ -0.0403 (0.100)
$\beta_{4,b,1}$ 0.0121 (0.900)	$\beta_{4,b,0}$ -0.0097 (0.901)	$\beta_{5,b,0}$ -0.0203 (0.199)	$\beta_{5,b,1}$ 0.0282 (0.065)*	$\beta_{6,b,0}$ -1.5894 (0.127)	$\beta_{6,b,1}$ -1.5667 (0.141)
$\gamma_{b,ind,0}$ 2.5558 (0.4302)	$\gamma_{b,ind,1}$ 2.1666 (0.0008)***	$\gamma_{b,us,0}$ 4.4891 (0.538)	$\gamma_{b,us,1}$ 2.1367 (0.394)	$\gamma_{b,ja,0}$ 1.0719 (0.636)	$\gamma_{b,ja,1}$ 0.4103 (0.883)
$\gamma_{b,uk,0}$ 1.5764 (0.815)	$\gamma_{b,uk,1}$ -3.8110 (0.221)	$\gamma_{b,ch,o}$ 0.0000 (1.000)	$\gamma_{b,ch,1}$ -0.0007 (0.998)	$\gamma_{b,em,0}$ 3.8111 (0.0006)***	$\gamma_{b,em,1}$ 5.6874 (0.0027)***
σ_s^2 0.0002 (0.0026)***	b_{ss} 0.7631 (0.0001)***	c_{ss} 0.1346 (0.0011)***	σ_b^2 0.0002 (0.0001)***	b_{bb} 0.1882 (0.104)	c_{bb} 0.3054 (0.0011)***
$\rho_{s,b}$ 0.1928 (0.0001)***					

Notes: ****, ** and * denote statistical significance at the 1 per cent and 10 per cent levels respectively.*

References

Baba, Y., R. Engle, D. Kraft, and K. Kroner, “Multivariate Simultaneous Generalized ARCH”,*Unpublished Manuscript, University of California at San Diego, Dept. of Economics,* 1989.

Bollerslev, T., “Modeling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized Approach”, Review of Economics and Statistics, 1990, Vol. 72, pp. 498-505.

Campbell, J., and J. Ammer, “What moves the Stock and Bond Markets? A Variance Decomposition for Long-term Asset Returns”, *Journal of Finance*, 1993, Vol. 48, pp. 3-37.

Fama, E., and K. French, “Business Conditions and Expected Returns on Stocks and Bonds”, *Journal of Financial Economics*, 1989, 25, pp. 23-49.

Fama, E., and K. French, “Efficient Capital Markets: II”, *Journal of Finance*, 1991, Vol. XLVI, no. 5, pp. 1575-1617.

Fama, E., and K. French, “Common Risk Factors in the Returns on Stocks and Bonds”, *Journal of Financial Economics*, 1993, 33, pp. 3-56.

Gallagher, L., and C. Twomey, “ Identifying the Source of Mean and Volatility Spillovers In Irish Equities: A Multivariate GARCH Analysis”, *Economic and Social Review*, 1998, Vol. 29, pp. 341-356.

Keim, D. and R. Stambaugh, “Predicting Returns in the Stock and Bond Markets”, *Journal of Financial Economics*, 1986, 17, pp. 357-390.

Phillips, P, and Perron, P., “Testing for a unit root in Time Series Regression”, *Biometrika*, 1988, Vol. 75, No. 2, pp. 335-346.

Raj, J. and S. Dhal, “Integration of India’s Stock Market with Global and Major Regional Markets”, *Bank of International Settlement Report, (ed.) Regional Financial Integration in Asia: Present and Future*, 2008, pp. 222-236.

Sharma, J., and R. Kennedy, “Comparative Analysis of Stock Price Behavior on the Bombay, London & New York Stock Exchanges”, *Journal of Financial and Quantitative Analysis*, 1977, pp. 391-403.

Stivers, C., and L. Sun, “Stock Market Uncertainty and the Relation between Stock and Bond Returns”, *Federal Reserve Bank of Atlanta Working paper,* 2002, no. 3, March 2003.

Wong, W., A. Agarwal, and J. Du, “Financial Integration for India Stock Market, A Fractional Cointegration Approach”, *Department of Economics Working Paper No. 0501, National University of Singapore*, 2005.

Worthington, A., and H. Higgs, “Transmission of Equity Returns and Volatility in Asian Developed and Emerging Markets: A Mulitvariate GARCH Analysis”, International Journal of Finance and Economics, 2004, Vol. 9, pp. 71-80.

FORECASTING USING FUZZY MULTIPLE OBJECTIVE LINEAR PROGRAMMING

Kenneth D. Lawrence, School of Management, New Jersey Institute of Technology, Newark, NJ 07102, carpetfour@yahoo.com, 1-973-596-6425

Dinesh R. Pai, Penn State Lehigh Valley, Center Valley, PA 18034, drp18@psu.edu, 1-610-285-5029

Sheila M. Lawrence, Department of MSIS, Rutgers University, Piscataway, NJ 08854, lawrencesm5@aol.com, 1-973-596-6425

ABSTRACT

The application of fuzzy approach models to forecast has been drawing a lot of attention. This study proposes a fuzzy approach to forecasting using a financial data set. The methodology used is multiple objective linear programming (MOLP). Selecting an individual forecast based upon a single objective may not make the best use of available information for a variety of reasons. Combined forecasts may provide a better fit with respect to a single objective than any individual forecast. We incorporate soft constraints into a mathematical programming approach to improve our forecasting accuracy. We compare the results of our approach with the preemptive MOLP approach. A financial example is used to illustrate the efficacy of the proposed forecasting methodology.

Keywords: MOLP, Linear programming, Fuzzy, Forecasting

INTRODUCTION

An important problem facing decision makers in business organizations is the forecasting of uncertain events. The importance of having accurate forecasts available for decision making is widely recognized at all levels of decision making and in all functional areas of business. Because of the importance of the problem, multiple forecasts are often prepared for the same event. Decision makers have to evaluate the multiple forecasts with respect to a single objective and select the forecast method which comes closest to satisfying the chosen objective, while the remaining forecasts are discarded. Selecting a single forecast based upon a single objective may not make the best use of available information for a variety of reasons. Although there is likely to be some overlap or redundancy of information among forecasts, discarded forecasts may contain information not available in the selected forecast. Combined forecasts may provide better fit with respect to a single objective than individual forecast. Even if an individual forecast does provide a good fit with respect to a single objective, a combined forecast may provide a better fit with respect to multiple objectives.

The purpose of this study is to investigate fuzzy approach to a combination of several techniques for forecasting monthly sales to produce improved forecasts over those produced by more traditional single technique approaches. This paper compares the forecast provided by combining the traditional forecasting with a fuzzy approach using soft constraints. The methodology used for combining the forecasts is preemptive multiple objective linear programming (MOLP).

LITERATURE REVIEW

Combining forecasts introduced by Bates and Granger (1969) is often considered as a successful alternative to using just an individual forecasting method. Empirical results demonstrate that no single forecasting method can generate the best forecasts in all situations and the relative accuracy of the different models varies with the origin/destination pairs and the lengths of the forecasting horizons [11]. Previous studies have shown that composite forecasting is useful in predicting variables of interest such as sales, corporate earnings per share, tourists inflow, etc [3] [10] [7] [12]. Wong et al. (2007) show that forecast combination can improve forecasting accuracy and considerably reduce the risk of forecasting failure. They conclude that combined forecasts can be preferred to single model forecasts in many practical situations. In a recent study, Hibon and Evgeniou (2005) propose a simple model-selection criterion to select among forecasts. Their results indicate that the advantage of combining forecasts is not only better results but also that it is less risky in practice to combine forecasts than to select an individual forecasting method.

The first application of forecasting using fuzzy set theory to our knowledge appeared in [4]. The author applied fuzzy concepts on the computer simulation for power demand forecasting and loading of power systems. Since this initial research, the Interest in fuzzy forecasting has grown considerably. Shnaider and Kandel (1989) develop a computerized forecasting system to forecast corporate income tax revenue for the state of Florida. Song and Chissom (1993) provide a theoretic framework for fuzzy time series modeling.

THE MODELS

Multiple Objectives Linear Programming

Preemptive multiple objective linear programming (MOLP) techniques are used to generate efficient combined forecasts. The forecasting techniques utilized in the study are exponential smoothing, multiple regression, and harmonic smoothing. A basic MOLP model for combining forecasts can be formulated as follows. The decision variables in the model are defined as the weights to be assigned to each forecast:

$W_j =$ The weight assigned to forecast j, $\quad j = 1, 2, ..., n$

The coefficients of the model are the actual observed values and the forecasted values for each of the forecasts in each of the time periods considered:

$A_i =$ The actual observed value in time period i, $\quad i = 1, 2, ..., m$

$F_{ij} =$ The forecasted value by forecast j in time period i, $\quad i = 1, 2, ..., m, \; j = 1, 2, ..., n$

The constraints of the model take the following form:

$$\sum_{j=1} F_{ij} W_j + d_i^- + d_i^+ = A_i \qquad i = 1, 2, ..., m \tag{1}$$

Where,

$d_i^- =$ The underachievement by the combined forecast of the observed value in time period i,

$d_i^+ =$ The overachievement by the combined forecast of the observed value in time period i,

and

$$\sum_{j=1}^{n} W_j = 1 \tag{2}$$

$$W_j \geq 0, \qquad j = 1, 2, ..., n$$

$$d_i^-, d_i^+ \geq 0, \qquad i = 1, 2, ..., m$$

In each time period, the weighted sum of the forecasted values plus or minus an error term must equal the actual observed value. The objectives of the model are expressed in terms of the underachievement and overachievement variables:

Minimize: $\quad Z = \quad [Z_1(\bar{d}^-, \bar{d}^+), \; Z_2(\bar{d}^-, \bar{d}^+), \; ..., \; Z_k(\bar{d}^-, \bar{d}^+)]$

Where $\quad \bar{d}^- = (d_1^-, d_2^-, ..., d_m^-), \quad \bar{d}^+ = (d_1^+, d_2^+, ..., d_m^+)$

The preemptive MOLP model objectives as formulated in (3) represent alternative measures of forecast error or accuracy. Alternative accuracy objectives allow decision makers to emphasize their preferences for the form and occurrence of forecast error. Two examples of such alternative measures are: (1) the minimization of total forecast error over all time periods, and (2) the minimization of the maximum forecast error in any individual time period. A third measure of error, which could be stated as an objective, is the minimization of forecast error in the most recent time periods. Thus, the overall MOLP model for combining forecasts consists of minimizing (3) subject to (1)-(2).

The associated objective functions are:

$$Z_1 = \sum_{i=1}^{30} (d_i^- + d_i^+)$$

$$Z_2 = \sum_{i=1}^{30} d_i^+$$

$$Z_3 = \sum_{i=25}^{30} (d_i^- + d_i^+)$$

Fuzzy Approach

The fuzzy goal programming (FGP) model enlarges the feasible region by fuzzifying the constraints of the model with given tolerance values. Here, we use combination soft and crisp constraints instead of only crisp constraints in the MOLP. The fuzzy approach formulation is as follows [5]:

Minimize: $Z = \quad [Z_1(\bar{d}^-, \bar{d}^+),\ Z_2(\bar{d}^-, \bar{d}^+),\ ...,\ Z_k(\bar{d}^-, \bar{d}^+)]$

Subject to : $\sum_{j=1} \tilde{F}_{ij} W_j + d_i^- + d_i^+ = \tilde{A}_i \qquad i = 1, 2, ..., m$

and $\sum_{j=1}^{n} W_j = 1, \quad W_j \geq 0, \quad j = 1, 2, ..., n$

$d_i^-, d_i^+ \geq 0, \ i = 1, 2, ..., m$

Where $\tilde{F}_{ij}, \tilde{A}_i$, $i = 1, 2, ..., m$, $j = 1, 2, ..., n$ are fuzzy coefficients in terms of fuzzy sets.

RESULTS

Z_1 Total Forecast Error	Z_2 Positive Forecast Error	Z_3 Recent Forecast Error	W_1 Exponential Smoothing	W_2 Harmonic Smoothing	W_3 Multiple Regression	Z_1 Total Forecast Error	Z_2 Positive Forecast Error	Z_3 Recent Forecast Error
414.89	232.82	123.04	0.60	0.09	0.31	108.25	108.25	44.01

Table 1. Results

Tables 1. Shows the results of our models. Several observations can be made based on the above results. First, combining different forecasting methods helps improve the level of one or more objectives without worsening the level of any other objective. Second, using similar weights as in MOLP, the preemptive fuzzy approach give remarkably improved results.

REFERENCES

[1] Bates, J. M., Granger, C. W. J. “The combination of forecasts.” *Operational Research quarterly*, 1969, 20, 451–468.

[2] Chen, S. M. “Forecasting enrollments based on fuzzy time series.” *Fuzzy Sets and Systems*, 1996, 81(3), 311-319.

[3] Conroy, R., Harris, R. “Consensus forecasts of corporate earnings: analyst’s forecasts and time series methods.” *Management Science*, 1987, 33,725-738.

[4] Economakos, E. Application of fuzzy concepts to power demand forecasting. *IEEE Transactions on Systems, Man and Cybernetics*, 1979, 9(10), 651-657.

[5] Fang, S., Hu, C., Wang, H., Wu, S. “Linear Programming with Fuzzy Coefficients in Constraints.” *Computers and Mathematics with Applications*, 1999, 37 63-76

[6] Hibon, M., Evgeniou, T. To combine or not to combine: selecting among forecasts and their combinations. *International Journal of Forecasting*, 2005, 21(1), 15-24.

[7] Huang, H., Lee, T. “To Combine Forecasts or to Combine Information?” No 200806, Working Papers from University of California at Riverside, Department of Economics, 2007.

[8] Shnaider, E., Kandel, A. “The use of fuzzy set theory for forecasting corporate tax revenues.” *Fuzzy Sets and Systems*, 1989, 31(2), 187-204.

[9] Song, Q., Chissom, B. “Fuzzy time series and its models.” *Fuzzy Sets and Systems*, 1993, 54(3), 269-277.

[10] Reeves, G. R., Lawrence, K.D. “Combining multiple forecasts given multiple 0bjectives.” *Journal of Forecasting*, 1982, 1, 271-279.

[11] Witt, S. F., Song, H. “Forecasting tourism flows.” In A. Lockwood, & S. Medlik (Eds.), *Tourism and hospitality in the 21st Century*, 106–118, 2002.

[12] Wong, K., Song, H., Witt, S., Wua, D. “Tourism forecasting: To combine or not to combine?” *Tourism Management*, 2007, 28(4), 1068-1078.

MULTIPLE OBJECTIVE SIMULATION MODELING TO DEVELOPING GREATER EFFICIENCY IN HEATH CARE SUPPLY CHAINS TREATING PATIENTS AS INVENTORY

Virginia M. Miori, Ph.D.
St. Joseph's University
Philadelphia, PA
vmiori@sju.edu

ABSTRACT

The authors have previously validated a design of the health care supply chain which treats patients as inventory without loss of respect for the patients. This work continues examination of patients as inventory while addressing the dual objectives of reducing redundancy in services and creating greater efficiency in the health care supply chain. Health care professionals have become increasingly specialized, resulting in much more flexible and more expensive supply chains. The lack of common data storage has created a need for redundancy (or rework) in medical testing. The overlap between these objectives is significant though the origins of the constraints are varied. We perform simulations using SigmaFlow software to address these objectives relative to the resource constraints, monetary constraints and the overall culture of the medical supply chain. The simulation outcomes lead us to recommendations for data warehousing as well as providing a mechanism such as inventory postponement strategies to establish appropriate lines between completely efficient and completely flexible health care supply chains.

Keywords: supply chain, simulation, SigmaFlow

INTRODUCTION

After having validated a model of the health care supply chain (HCSC), treating patients as inventory, we now approach two of the most pressing issues faced within the HCSC. In essence, the patient comes to us as a subassembly awaiting the use of other resources. By adapting this view, we may continue to evaluate supplies, medical professionals and medications as resources. In addition, we allow a different perspective which facilitates the use of alternative modeling approaches, and therefore develop the potential for increasingly effective practical solutions.

The approach is simple to state; a patient will be viewed as inventory. Yet, the implications of this statement are quite complex. The concept of humans as inventory has been discussed in a number of divergent areas of research. Industrial engineers refer to it in time and motion studies. Schools refer to it in terms of student tracking. Project managers refer to it in resource utilization, and human resource departments refer to it in staffing models. Let us not forget the greatest use of the concept, throughout history human inventory has been taken in the form of censuses. All of these areas conjure in our minds the same ethical dilemma; they pose the possibility of dehumanizing individuals.

A thorough discussion of the ethical implications may be found in Miori (2009). We move forward through this research with care and endeavor to always maintain a respectful view of our patients. We hold their interests in the highest regard and we hold ourselves responsible for compassionate treatment of the subject, resulting in compassionate and effective treatment of the patients.

Among the many issues faced by the HCSC, redundancy in testing and treatments is of primary concern. Due to the lack of common data storage for patient treatment records, circumstances often lead to repeating the same tests in order to establish appropriate treatment regimes. In some cases this may also lead to redundancy of medications and potentially unsafe conditions for the patient.

Patients may visit multiple doctors for various reasons such as travel, lack of availability of the same physician, dissatisfaction, second opinions, or even a desire to exceed therapeutic doses of certain medications. The additional health care costs along with the potential risks to patients support our emphasis on reduction of data redundancy.

The second critical issue that we seek to address is that of HCSC efficiency. By its nature, the HCSC is a very flexible supply chain. Individual patient needs are fully assessed before the development of a customized treatment plan. Though it is important to meet all patient care needs, we explore the possibility of organizing the patient treatment process in a way that can yield more efficient and timely treatment of patients. Attempts to achieve this have been made by many hospitals and hospital systems. The results have ranged from the creation of flex care units within emergency rooms to certain hospitals specializing in types of care.

In this paper, we consider the multiple objectives and propose improvements to the supply chain. The improvements are borne out of inventory postponement strategies, and secure and efficient data warehouse techniques.

The background for this paper covers many areas. We first present related literature in these many areas. The current health care supply chain is discussed with the presentation of a simulation model and baseline results. The proposed inventory interpretation and potential inventory models are shown. The revised supply chain is presented with modified inventory modeling with a revised simulation model. Preliminary results and recommendation are presented along with future research directions.

LITERATURE

Available upon request.

HCSC

Service and production supply chains are a direct reflection of the goals and strategic objectives of the organization. In a medical setting, a supply chain decision to be made is the selection of an area of specialization. It has become common, especially in larger cities, for hospitals to focus on particular areas of medical care. These areas include general medicine, heart health, trauma care, cancer treatment, neck and spinal care, pediatric care as well as many others. In the case of this research, the focus of the hospital is general medicine.

The patient-as-inventory HCSC focuses on the processes that must be performed to yield a finished product. The focus during the processes is on an appropriate conversion of the input (the ailing patient) to the output (the released patient).

The general HCSC is shown in Figure 1. The point of contact in the community acts as a supplier, providing ailing patients to the system. The medical care includes what we consider to be production processes and the final product is the discharged patient.

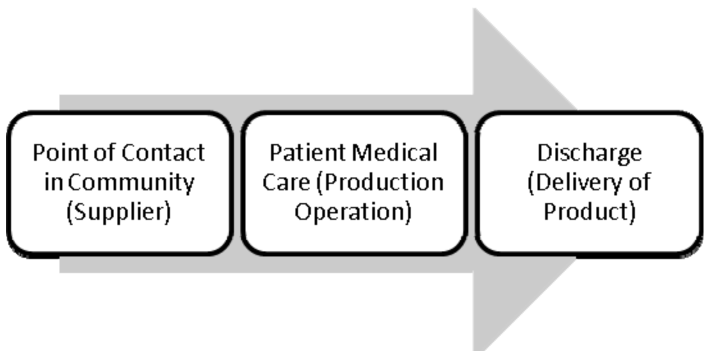


Figure 1: Health Care Supply Chain

To more thoroughly describe the health care supply chain (HCSC), we begin by visiting it from several perspectives. Supply chains are often described as some combination of push and pull along with the push/pull boundary. They are also discussed with reference to their efficiency and responsiveness. Both of these methods are fitting in the HCSC setting.

Companies are considered to be efficient if their products or services are standardized. There is little if any variation in the end product and therefore the product or service is highly repeatable and highly reproducible. Any number of people can be trained to affect the same result in the same way. Consumer products companies fall into this category as well as clothing staples such as underwear. In the service sector, fast food restaurants fall into this category.

The antithesis of the efficient company is the responsive company. Responsive companies deal in high levels of specialization and employees tend to be well trained in a single area. The resulting products or services are generally unique to a particular customer. Health care requires a great deal of specialization and is a highly responsive industry. Our supply chain, therefore, is characterized as a responsive supply chain as shown in Figure 2.

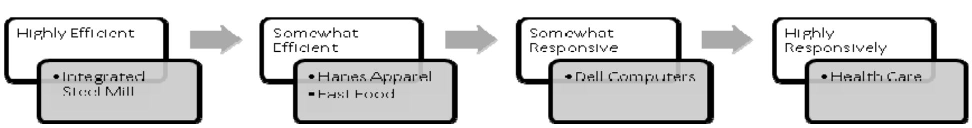


Figure 2: Efficient versus Responsive Supply Chains

The discussion of efficiency versus responsiveness provides an excellent segue into the examination of the push/pull boundary within the HCSC. An entirely push process is one that produces products to inventory. An entirely pull process is one that completely produces products to specific product design needs. The push/pull boundary provides the point at which any process converts from push to pull.

In the HCSC setting, we are tempted to say that it is an entirely pull process, but we must take note of common functions performed on all patients. Much to our dismay, the first step is completion of paperwork, no matter what condition a patient is in. Coincident with the paperwork, triage is performed on every patient to assess their needs. At this point, the HCSC becomes a pull process. All treatments are dictated by the condition of the individual patient. We first expand the diagram presented in Figure 1, and provide a cycle view of the HCSC in Figure 3. In the cycle view, we begin to break down aspects of the care into logical cycles. We then use these cycles to describe the push pull boundary in Figure 4.

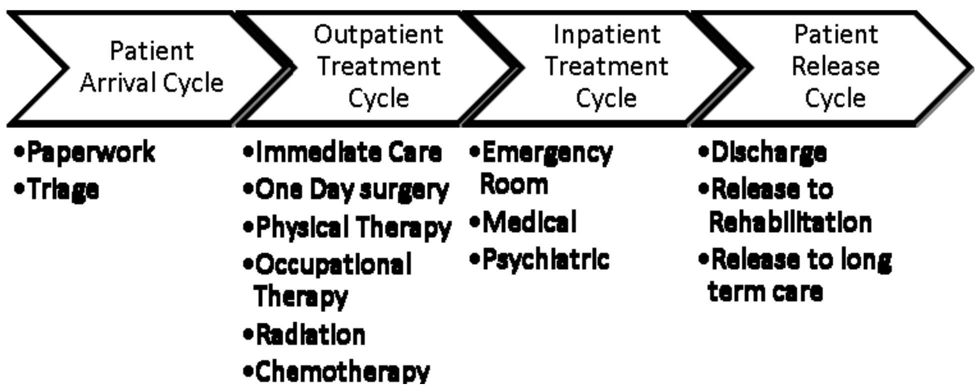


Figure 3: Cycle View of Health Care Supply Chain

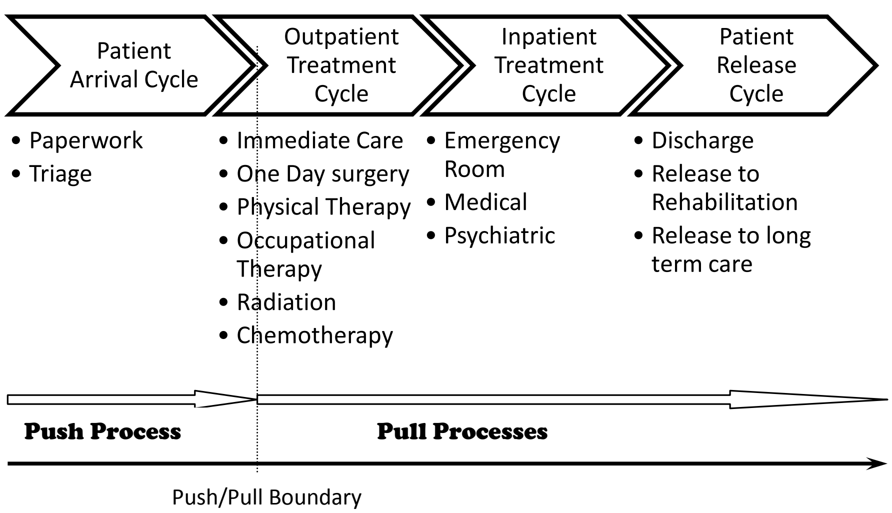


Figure 4: Push/Pull View of Supply Chain Processes

Through the characterization of the supply chain as responsive and heavily pull process, we exemplified the goals of the health care provider.

SIMULATION OF HCHS

The simulations were performed using SigmaFlow simulation software. SigmaFlow simulations are created with a detailed process maps augmented with parameters describing arrivals, processes and departures. The first map, shown in Figure 5 (available upon request), shows the current state of the supply chain as the model benchmark. The horizontal bands within the process map allow us to categorize those responsible for treatment and/or care. Each procedure block is characterizes by the expected duration of stay in that area and the possible paths through other areas.

The individual departure paths from each procedure block are assigned a percentage which indicates the likelihood of each successive procedure block. The duration of each procedure is designated as a statistical distribution. In addition to the model benchmark, two other scenarios were run. The first represented the improvements that would result from establishment of a central regional data warehouse and the second represented the improvements expected from efficiencies gained in patient evaluation and treatment.

The issue of redundancy in testing is especially problematic when looking at the duration of emergency room (ER) care. Currently medical records are electronically available on a limited basis, and then information is typically restricted to doctors within the particular practices treating individual patients. This lack of availability limits an ER's ability to obtain accurate medical histories, particularly in the case of chronic or preexisting conditions. An ER has no choice but to re-administer any tests they deem necessary. Introduction of a central regional data warehouse is therefore critical to reductions in the time required for ER care. We estimate that the average time consumed by each patient during check-in and evaluation could be reduced by 30% with the introduction of the central regional data warehouse.

Flexibility in health care supply chains has been a significant goal in the continuous quest for better medical care. Unfortunately it has also been one of the catalysts for medical expenses spiraling out of control. The sweeping health care reform bills that have been carried through congress are clear signs of our need to reduce the cost of illness. As we noted, efficiency is the counter position to flexibility in supply chains. By shifting our HCSC on this continuum to a position of greater efficiency, we may again decrease time required for treatments and generate more effective resource utilization. Our chief strategy to accomplish this is inventory postponement.

Recall that in a manufacturing setting, we may describe inventory postponement as the maintenance of a standard set of processes to be performed, thus delaying specialized processes until the last possible moment. In doing so, these standard procedures allow the maintenance of efficient processes as long as possible before customizing a product for delivery. In this paper, with patients are our inventory; we therefore seek to extend the current standard processes into a more complete serried of diagnostic procedures. In conjunction with available medical histories, this would allow an ER to develop a thorough patient health baseline.

Upon completing the baseline simulation, the first scenario will reflect the time savings accrued from centralized data warehousing. The second scenario will reflect potential efficiencies that can result from our strategy of inventory postponement. We estimate that the average time consumed by each patient during management could be reduced by 20% with the introduction of greater efficiencies.

The preliminary results of the first simulation scenarios not only show the expected processing time improvements, but they also show considerable improvements in waiting time for patients. The results are presented in Table 1. Scenario 2 did show improvements in processing time but did not show any appreciable improvements in queue time.

Table 1: Simulation Results

	% Improvement Scenario 1 over Baseline	% Improvement Scenario 2 over Baseline
Average Work Time	30.3%	19.8%
Average Queue Length	25.6%	0
Maximum Queue Length	25.0%	0
Average Queue Time	25.5%	0
Maximum Queue Time	56.6%	0

CONCLUSIONS

This work shows that the use of a central regional data warehouse would significantly improve efficiency within the medical supply chain by making more patient information available. Redundancy in testing would be avoided along with elimination of inaccuracies in patient histories. The use of inventory postponement would not accrue any improvements in waiting time for patients despite reductions in treatment times. The bulk of the waiting time experienced by patients occurs within the check-in and evaluation stages of emergency room care. Through there were no improvements in waiting time, inventory postponement practices are still helpful in managing resources.

References available upon request.

DATA ENVELOPMENT ANALYSIS (DEA) EFFICIENT LOCATION/ALLOCATION DECISION MODELS II

Ronald K. Klimberg, Haub School of Business, Saint Joseph's University, Philadelphia, PA, 19131, 610-660-1625, klimberg@sju.edu.

Samuel J. Ratick, The George Perkins Marsh Research Institute, The Graduate School of Geography, Clark University, Worcester, MA 01610, SRatick@clarku.edu.

Vinay Tavva, Saint Joseph's University, Haub School of Business, Philadelphia, PA, 19131, 610-660-1625, vt441328@sju.edu.

Sasanka Vuyyuru, Saint Joseph's University, Haub School of Business, Philadelphia, PA, 19131, 610-660-1625, sv464020@sju.edu.

Daniel Mrazik Saint Joseph's University, Haub School of Business, Philadelphia, PA, 19131, 610-660-1625, daniel.mrazik@sju.edu.

ABSTRACT

There have been many facility location/allocation models developed to find optimal spatial patterns with regard to various location criteria including: cost, time, coverage, and access, among others. In a previous study we developed and tested location modeling formulations that utilize data envelopment analysis (DEA) efficiency measures to find optimal and efficient facility location/allocation patterns, [6]. These models solve the DEA efficiency measure and the location modeling objectives simultaneously. In this paper, we extended our capacitated adjustable simultaneous DEA/CPLP formulation (CASD) model, to allow the outputs to be a variable. In such a case, the model is now nonlinear. We solve this nonlinear model and apply it to a small hypothetical data set and assess the results.

INTRODUCTION

The initial focus of mathematical programming models for facility location was based upon the spatial availability of a server to fulfill demand (for a review of covering models see [7] and [9]. In [13] Toregas et al. formulated the Location Set Covering Problem (LSCP) that required coverage of all demand points within some time standard. A derivative of this was the *p*-center problem, which minimized the maximum distance separating any demand point from its nearest facility, [4] and [8]. Both of these models considered demand, server capability, and service time to be homogeneous. Budgetary constraints, which might make full coverage not possible, led to the development of the Maximal Covering Location Problem (MCLP), [2] and [14]. Non-uniform distribution of demand nodes was incorporated into these partial covering models to develop the best deployment of a given number of facilities.

In [3], Current, Min, and Schilling review of 45 facility location papers demonstrated that these problems are inherently multiple objective. They classified the most common objectives into four categories: cost minimization, demand-oriented, profit maximization and, environmental concern. The cost minimization objective is the traditional objective of most facility location models. Demand-oriented objectives focus on measuring the "closeness" of the facilities, where "closeness" may be measured in terms of coverage or response time. A good location pattern is one that not only optimizes the spatial interaction among facilities, and the demands they serve, such as the above mentioned models, but additionally, optimizes the performance (efficiency) of those facilities at the chosen locations. To evaluate the efficiency of the facilities we use a linear programming approach called data envelopment analysis (DEA).

DEA, which has been used for about thirty years, was developed to address the common problem of measuring organizational efficiency when many measures of inputs and outputs are available, [1]. DEA utilizes linear programming to produce measures of the relative efficiency of comparable units that employ multiple inputs and outputs. It requires only that the selected inputs and outputs be quantifiable. The technique can analyze these multiple inputs and outputs in their natural physical units without reducing or transforming them into some common measurement such as dollars. DEA takes into account these multiple inputs and outputs to produce a single aggregate measure of the relative efficiency of each comparable unit. DEA defines efficiency as the ratio of weighted outputs to weighted inputs:

$$\text{Efficiency} = \frac{\text{weighted sum of outputs}}{\text{weighted sum of inputs}}\ .$$

The more output produced for a given amount of resources the more efficient (i.e., less wasteful) is the operation. DEA uniquely evaluates all the DMUs and all their inputs and outputs simultaneously, and conservatively identifies the sets of relatively efficient and relatively inefficient DMUs. Thus, the solution of a DEA model provides a manager a summary with comparable DMUs grouped together and ranked by relative efficiency. DEA has been applied to several diverse areas as hospitals, health care organizations, physicians, pharmacies, drug reimbursement, armed forces, criminal courts, schools, university departments, banks, electric utilities, strip mining, manufacturing productivity, and railroad property evaluation, [5], [10], [11], and [12].

Typically solving a DEA model consists of subsequently solving a slightly modified linear program for each DMU. However, to link DEA to a facility location model, the DEA model should be modified such that the DEA efficiencies of all the DMUs are calculated in one linear program. Unlike typical DEA in which the efficiency of each DMU is maximized, we developed a model (SDEA) in Klimberg & Ratick to solving DEA for each DMU simultaneously. In this case, this simultaneous DEA model maximizes the sum of the efficiencies for all of the DMUs. In [6], we associated this simultaneous DEA model with to the capacitated facility location model (CPLP). In the case of the capacitated model, we assigned one DEA output variable to the amount demand. As a result, the DEA efficiencies are now dynamic, i.e., the amount of output may change depending on how much is supplied and hence the DEA efficiencies will change. An addition aspect considered is that all potential locations are not necessarily opened while DEA considers the efficiencies of all locations. In our formulation, the efficiency of a location is equal to 0 if it is not used. The combination of the SDEA model with the CPLP, results in a capacitated adjustable simultaneous DEA/CPLP formulation (CASD), [6].

In order to keep the CASD formulation linear, we allowed a fraction of one of the output variables to satisfied demand. In such a case, the fractional contribution of that output to the

DEA score is not properly weighted in the DEA solution. If we keep the true DEA weight and allowed the amount of output to vary as given in the CASD formulation, it would have resulted in a non-linear program. In this paper, we extend the CASD formulation to address this issue by allowing all the output variables to varying—changing them from parameters to variables, and hence, allowing the DEA scores to dynamically change.

The formulation for this capacitated adjustable simultaneous dynamic DEA/CPLP model ($CASD^2$) is the following:

Model: Capacitated Adjustable Simultaneous Dynamic DEA/CPLP model ($CASD^2$)

$$\text{MAX} \sum_{k=1}^{K} \sum_{l=1}^{L} (1 - d_{kl}) \quad (1)$$

$$\text{MIN} \sum_{k=1}^{K} \sum_{l=1}^{L} c_{kl} b_{kl} + \sum_{k=1}^{K} F_k y_k \quad (2)$$

s.t.

$$\sum_{k=1}^{K} x_{kl} \geq 1 \quad \forall l \quad (3)$$

$$x_{kl} \leq y_k \quad \forall k, l \quad (4)$$

$$\sum_{k=1}^{K} b_{kl} = \text{dem}_l \quad \forall l \quad (5)$$

$$\sum_{j=1}^{J} p_{jkl} = b_{kl} \quad \forall k,l \quad (6)$$

$$p_{jkl} \leq O_{jkl} y_k \quad \forall k,l \quad (7)$$

$$\sum_{i=1}^{I} v_{kli} I_{kli} = x_{kl} \quad \forall k, l \quad (8)$$

$$\sum_{j=1}^{J} u_{klj} p_{jkl} + d_{kl} = x_{kl} \quad \forall k, l \quad (9)$$

$$\sum_{j=1}^{J} u_{klj} p_{jrs} - \sum_{i=1}^{I} v_{kli} I_{irs} \leq 0 \quad \forall k,l; \forall r, s; (k \neq r \text{ and } l \neq s) \quad (10)$$

$$u_{klj} \geq \varepsilon x_{kl} \quad \forall k,l,j \quad (11)$$

$$v_{kli} \geq \varepsilon x_{kl} \quad \forall k,l,i \quad (12)$$

$$u_{klj} O_{jkl} \leq x_{kl} \quad \forall k,l \quad (13)$$

$$b_{kl} \geq x_{kl} \quad \forall k,l \quad (14)$$

$$y_k, x_{kl} = 0,1$$

$$b_{kl}, u_{klj}, v_{kli}, p_{jkl} \geq 0$$

where

$i = 1, \ldots, I$ inputs used at facility location

$j = 1, . . ., J$ outputs produced at facility location
$k = 1, . ., r, . ., K$ facility locations
$l = 1, . ., L$ index of demand locations

Parameters:

c_{kl} = cost of shipping one unit of demand from facility k to demand l
F_k = fixed cost of opening/using facility k
O_{jk} = amount of the jth output for the kth facility location
I_{ik} = amount of the ith input for the kth facility location

Decision Variables:

$$x_{kl} = \begin{cases} 1 & \text{if facility } k \text{ serves demand } l \\ 0 & \text{o/w} \end{cases}$$

$$y_k = \begin{cases} 1 & \text{if facility } k \text{ "opened" (i.e. used)} \\ 0 & \text{o/w} \end{cases}$$

b_{kl} = # of units shipped from facility k to demand location l
p_{jkl} = # of units of output j shipped from facility k to demand location l
u_j = the weight assigned to the jth output;
v_i = the weight assigned to the ith input.

The first objective function in the above formulation, (1) and the corresponding DEA-related constraints (8 – 10) are similar to and are related to the DEA portion of the model. The main difference being the additional index and the value x_{kl}. If facility k serves demand l, i.e., x_{kl} =1, the corresponding input and output weights are required to be greater than ε because of constraints (11) and (12) respectively; the constraints in (13) require the weighted outputs to be less than 1, for all facilities, demands, and output types. On the other hand, if facility k does not serve demand l ($x_{kl} = 0$), the constraints in (11) an (12) require the input and output weights to be non-negative, and the constraints in (8) and (13) force them to be equal to 0. The constraints in (14) require at least one unit to be shipped from facility k to demand l, if $x_{kl} = 1$, allowing for that facility to be used in the calculation of DEA efficiency scores. Thus, if x_{kl} =0, the DEA input and output weights for facility k are equal to 0, and that facility is not considered in the computation of the relative DEA efficiency scores for that location/allocation pattern. The second objective function (2) and constraints (3 – 6) are related to the facility location part of the model. The second objective function (2) calculates the total cost (transportation and fixed opening costs) of supplying the demand in the system. The transportation costs is calculated as the product of the per unit transportation costs and the amount shipped from facility k to demand l, i.e., b_{kl}. Total fixed costs of opening facilities is obtained by summing over all facilities the product of: the fixed costs (F_k) and the integer variable, y_k, which is 1 if that facility is chosen to be opened in the optimal solution. The constraints represented in (3) assure that every demand is satisfied, and the constraints in (4) assure that only open facilities can supply demands. Constraint (5) assures that all demands are satisfied by shipments from open facilities. Constraint (6) requires that all the output supplied to demand location l must satisfy its demand. Constraint (7) forces p_{jkl} to equal zero if facility k is not opened and further limits p_{jkl} to be less than O_{jkl} is facility k is opened. In constraints (9) and (10), the parameter O_{jkl} is replaced by the decision variable p_{jkl}, thus, making these constraints nonlinear.

To solve the CASD^2 model, which is a non-linear with mixed integer (MINLP) problem, we used the AIMMS linear programming software and in particular their outer approximation algorithm (AOA). The AOA algorithm controls the interaction between a master MIP model (which we used the CPLEX solver), and the NLP submodel (which we used the SNOPT solver). On a 2.4GHz Intel Core 2 Duo T7700 with 3GB of RAM laptop it took about 10 – 20 minutes to run the model.

Using the example we created in [6], a series of multiobjective solutions for the CASD^2 model were obtained by varying the relative weights on the minimum efficiency objective. Since we are now allowing all three outputs to satisfy demand and as a result there are so much excess outputs such that demand can be easily satisfied while reallocating the outputs in such a way that initially each DEA score was 100%. Therefore, we multiplied each demand location by 10 to seek to diminish this situation. The multiobjective solutions to the CASD^2 model with the increased demands are listed in Table 1. More tradeoff between costs and DEA efficiency is illustrated in these solutions than the CASD model results in [6].

Table 1. Solution Values for the CASD^2 model with the demands increase (multiplied by 10).

Objective Function Values Total Fixed Costs	$2,991	$3,026	$3,113	$3,163	$3,200
Total Transport Costs	$7,786	$7,786	$7,786	$7,786	$7,786
Total Costs	$10,777	$10,812	$10,899	$10,949	$10,986
Number of Open Facilities	5	5	5	5	5
Number of facility-demand links	32	32	35	36	37
Total Sum of Efficiency Scores	30.80	30.88	34.74	35.99	37.00
Min Efficiency Score for Solution	0.743	0.757	0.912	1.00	1.00

CONCLUSIONS

In this paper, we extended our capacitated adjustable simultaneous (CASD) model we previously developed in [6] to consider the more realistic situation where the outputs are decision variables. In such a case, the DEA scores are allowed to dynamically change and the model becomes nonlinear. The solutions further demonstrate the interactions between different location patterns and the effects on facility efficiency performance.

The CASD^2 model provided a more realistic situation by allowing the outputs to vary. However, several issues still remain to be addressed. First, even though all the outputs were allowed to change, the outputs are likely related to different demands. As a result, the model should be expanded to address the demand for different outputs. Next, besides varying outputs the inputs should also be permitted to vary. In such a case, the outputs and inputs would in all likelihood change proportionally. Lastly, each of the nonlinear solutions the solutions were locally optimal. Other solution procedures, such as genetic algorithm and tabu search, which have been shown to be proficient in solving nonlinear and integer problems, should be examined. The results from the CASD^2 model, developed in this paper, continue to demonstrate that addressing the location's efficiency as part of the location selection problem provides perceptive information for decision makers.

REFERENCES

[1] Charnes A., Cooper, W.W. and E. Rhodes, "Measuring Efficiency of Decision Making Units," *European Journal of Operational Research.* 2, pp. 429-444, 1978.

[2] Church, R.L. and C. ReVelle, "The Maximal Covering Location Problems," *Papers of the Regional Science Association*, Vol. 32, No. 101, pp. 101-1188, Fall 1974.

[3] Current, J., Min, H. and D. Schilling, "Multiobjective analysis of facility location decisions," *European Journal of Operational Research*, 49, pp. 295-307, 1990.

[4] Hakimi, S. L., "Optimal Location of Switching Centers and the Absolute Centers and Medians of a Graph," *Operations Research*, 12, pp. 450-459, 1964.

[5] Klimberg, R.K. and D. Kern, "Understanding Data Envelopment Analysis (DEA)," *Boston University School of Management Working Paper*, pp. 92-44, 1992.

[6] Klimberg, R. K. and S. J. Ratick, "Modeling Data Envelopment Analysis (DEA) Efficient Location/Allocation Decisions, "*Computers & Operations Research*, 35, pp. 457-474, 2008.

[7] ReVelle, Charles, "Urban Public Facility Location," in *Handbook of Regional and Urban Economics,* Vol. II, edited by E.S. Mills, pp. 1053-1096, 1987.

[8] ReVelle, Charles and R. Swain, "Central Facilities Location," *Geographical Analysis*, 2, pp. 30-42, 1970.

[9] Schilling, D. A., V. Jayaraman, and R. Barkhi, "A Review of Covering Models in Facility Location," *Location Science*, 1, pp. 25-55, 1993.

[10] Seiford, L.M., "Data Envelopment Analysis: The Evaluation of the State of the Art (1978-1995)," *The Journal of Productivity Analysis*, 9, pp.99-137,1996.

[11] Seiford, L. M. and R. M. Thrall "Recent developments in DEA: the mathematical programming approach to frontier analysis." *Journal of Econometric,* 46: 7-38, 1990.

[12] Tavares, G., "A bibliography of Data Envelopment Analysis (1978-2001)," RUTCOR, Rutgers University, 2002; also available at http://rutcor.rutgers.edu/pub/rrr/reports2002.1_2002.pdf

[13] Toregas, C., R. Swain, ReVelle, C. and L. Bergman, "The Location of Emergency Service Facilities," *Operations Research,* 19:6, pp. 1363-1373, Oct 1971.

[14] White, J. and K. Case, "On covering problems and the central facility location problem," *Geographical Analysis,* Vol. 6, No. 3, 1974.

A GENETIC ALGORITHM APPROACH TO END-OF-LIFE DISASSEMBLY SEQUENCING FOR ROBOTIC DISASSEMBLY

Ahmed ElSayed[*], *Elif Kongar*[†] *and Surendra M. Gupta*[‡]

ABSTRACT

End-of-life (EOL) processing options include reuse, remanufacturing, recycling and proper disposal. In almost all cases, a certain level of disassembly may be required due to possible changes in the original product structure. Thus, finding an optimal or near optimal disassembly sequence is crucial to increasing the efficiency of the process. Disassembly operations are labor intensive, can be costly, have unique characteristics and cannot be considered as reverse of assembly operations. Since the complexity of determining the best disassembly sequence increases as the number of parts of the product grow, an efficient methodology is reuired for disassembly sequencing. In this paper, we present a Genetic Algorithm for disassembly sequencing of EOL products. A numerical example is provided to demonstrate the functionality of the algorithm.

Keywords: Disassembly sequencing, genetic algorithm, product recovery, robotics.

1. INTRODUCTION

The growing amount of waste generated by the end-of-life (EOL) of products has become a severe problem in many countries. This fact, couple with the decreasing number of landfills and virgin resources has led to extended product and/or producer responsibility policies that mandate manufacturers to take-back their products at the end of their lives. In response, the manufacturers are now seeking solutions to address the potential accumulation of large inventories consisting of technologically invalid and/or non-functioning products. The challenge is to process these products in an environmentally benign and cost effective manner. There are various ways to accomplish this task under the general umbrella of end-of-life (EOL) processing, including reuse, recycle, and storage for future use or proper disposal.

In majority of EOL processing, a certain level of disassembly may be required. Disassembly can be *partial* or *complete* and may use a methodology that is *destructive* or *non-destructive*. In this paper, the case of complete disassembly is considered where the components are retrieved by either destructive or non-destructive methodology.

[*] Ahmed ElSayed, Department of Computer Science and Engineering, University of Bridgeport, 221 University Avenue, School of Engineering, 141 Technology Building, Bridgeport, CT 06604, E-mail : aelsayed@bridgeport.edu.
[†] Elif Kongar, Ph.D., Departments of Mechanical Engineering and Technology Management, University of Bridgeport, 221 University Avenue, School of Engineering, 141 Technology Building, Bridgeport, CT 06604, USA, Phone: (203) 576-4379, Fax: (203) 576-4750, E-mail : kongar@bridgeport.edu.
[‡] Surendra M. Gupta, Ph.D., P.E., Mechanical and Industrial Engineering and Director of Laboratory for Responsible Manufacturing, 334 SN, Department of MIE, Northeastern University, 360 Huntington Avenue, Boston, MA 02115, U.S.A., Phone: (617) 373-4846, Fax: (617) 373-2921, E-mail : gupta@neu.edu.

When the electronic products are concerned, which is the focus of this paper, modeling EOL processing problems can be complex due to the number of components in the product structure. In these cases, exhaustive search algorithms may mathematically be inefficient making heuristic methods more appealing for obtaining near-optimal or optimal solutions. With this motivation, this paper presents a genetic algorithm for optimal or near optimal disassembly sequencing in the presence of constraints and precedence relationships.

2. LITERATURE REVIEW

In recent years, genetic algorithm (GA) has been gaining popularity for solving combinatorial and *NP*-complete problems. One of the multi-objective optimization applications was proposed by Valenzuela-Rendón and Uresti-Charre [1]. Keung *et al*. [2] also applied a multi-objective GA approach to a tool selection model. Lazzerini and Marcelloni [3] used GA in scheduling assembly processes. In the area of disassembly, Kongar and Gupta [4] proposed a GA for disassembly sequencing problem, while McGovern and Gupta [5] applied genetic algorithm to disassembly line balancing. For further literature on disassembly scheduling and processing, see Ilgin and Gupta [6], Gungor and Gupta [7], Lee *et al*. [8], and Lambert and Gupta [9].

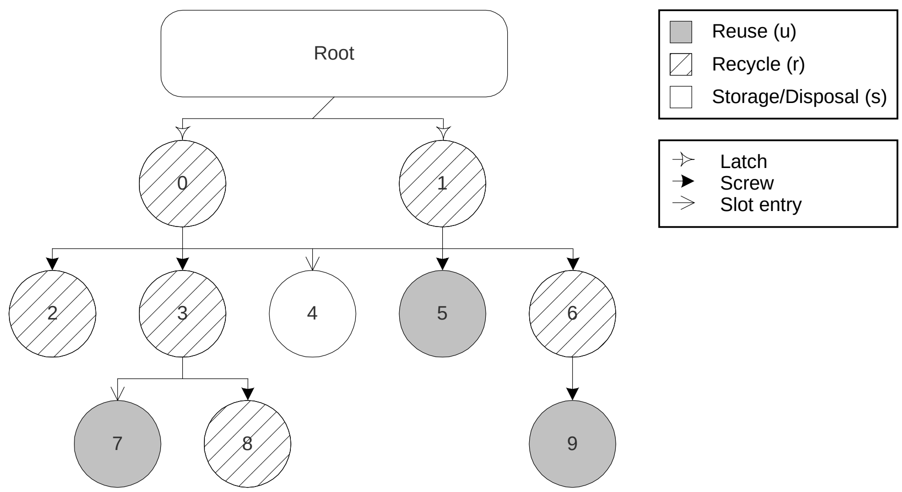
Precedence relationships are one of the factors that add to the complications in sequencing problems. In this regard, Sanderson *et al.* [10] considered precedence relationships in assembly sequence planning in such a manner. Seo *et al*. [11] proposed a genetic algorithm for generating optimal disassembly sequences considering both economical and environmental factors. Bierwirth *et al.* [12] and Bierwirth and Mattfeld [13] proposed a methodology to overcome this problem by introducing the precedence preservative crossover (PPX) technique for scheduling problems which preserves the precedence relationships during the crossover function of GA. This approach is also employed in this paper.

Furthermore, Shimizu *et al.* [14] developed a prototype system for strategic decision-making on disassembly for recycling at the design stage of the product life cycle. Hui *et al.* [15] utilized a genetic algorithm to determine feasible and optimal disassembly solutions.

In the area of automated disassembly, Torres *et al.* [16, 17] presented a personal computer disassembly cell that is able to handle a certain degree of automatism for the non-destructive disassembly process. This work is then followed by Pomares *et al.* [18] who generated an object-oriented model required for developing a disassembly process. Gil *et al.* [19] proposed a flexible multi-sensorial system for automatic disassembly using cooperative robots. As a follow-up work, Torres *et al.* [20] presented a task planner for disassembly process based on decision trees.

3. GENETIC ALGORITHM FOR DISASSEMBLY SEQUENCING

Genetic algorithm starts with a set of randomly selected potential solutions called the *population*. Each member of the population is encoded as an artificial *chromosome* which contains information about the solution mapping. Each chromosome is assigned a score based on a predefined *fitness function*. A new population of chromosomes is iteratively created in the hope of finding a chromosome with a better score. At each step of creation, a *mutation* may occur in a chromosome, and/or two chromosomes may mate to produce a child; this process is known as *crossover*. The selection of parent chromosomes is biased towards fitter members of the population. The process is iterated until some predetermined termination conditions are satisfied.



Item#	Description	Material	Disassembly		
			Time (sec)	Method	Coordinates (cm)
0	Back/Front Cover	Aluminium (A)	2	Destructive (D)	(30, 0,10)
1	Side Cover	Aluminium (A)	3	Destructive (D)	(30,40,10)
2	Power Supply	Copper (C)	3	Destructive (D)	(50,35, 5)
3	System Fan	Plastic (P)	2	Destructive (D)	(50,40, 5)
4	RAM	Plastic (P)	3	Non-Destructive (N)	(40,10, 0)
5	DVD Drive	Aluminium (A)	4	Non-Destructive (N)	(40, 7, 7)
6	Hard Drive Slot	Plastic (P)	2	Destructive (D)	(35, 5, 7)
7	CPU	Plastic (P)	1	Non-Destructive (N)	(40,20, 0)
8	Heat Sink Fan	Aluminium (A)	3	Destructive (D)	(35,20, 0)
9	Hard Drive	Aluminium (A)	2	Non-Destructive (N)	(40, 5, 7)

Figure 1. Original product structure of the EOL product

3. 2 Elements of Genetic Algorithm for Disassembly Sequencing

The structure of the example EOL product and related data are provided in Figure-1. The given product consists of ten components indexed by integers from 0 to 9. Therefore, $j \in \{0,1, .. , n\text{-}1\}$. Locations of each component is defined by their 3-D coordinates. The methodology used for the disassembly of each component may be destructive (*D*) or non-destructive (*N*). A component may or may not be demanded. If it is not demanded, it is represented by *s*. If it is demanded, it may be demanded for reuse (*u*) or recycling (*r*). A component has one of three types of joints, viz., latch (*L*), screw (*S*) or slot entry (*E*). In addition, the precedence relationships are given as follows: component 1 **or** 2 must be disassembled prior to any other component; component 3 must be disassembled prior to components 7 **and** 8; and component 6 must be disassembled prior to component 9.

Chromosomes: In every GA problem, the solution and parameters must be coded into chromosomes, represented by a combination of numbers, alphabets and/or other characters, before they can be processed. In this study, in order to capture the five variables, chromosomes are codified in the form of a string consisting of five ordered sections of equal length, representing the disassembly sequence, the disassembly method, the demand type of each component, and the material type of each component respectively. Coordinate data are kept separate from the chromosome structure to provide ease in calculations.

Initial Population: The initial population consists of *ncr* random chromosomes that satistify the precedence relationships and any other constraints imposed by the product structure. For the example provided in Figure-1, hundred chromosomes (feasible solutions) are randomly created to form the initial population (*ncr* = 100).

The chromosomes in the initial random population is provided in Table-1 (10 out of 100 are shown).

Table 1. Partial initial population for the genetic algorithm example

Sequence	Method	Demand	Material	$F(ch,gn)$
0623795418	DDDDNNNNDD	rrrruuusrr	APCPPAAPAA	28.978
0613275948	DDDDDNNNND	rrrrruuusr	APAPCPAAPA	29.881
1630824957	DDDDDDNNNN	rrrrrrsuuu	APPAACPAAP	30.563
0327468195	DDDNNDDDNN	rrrusrrruu	APCPPPAAAA	30.586
1328047659	DDDDDNNDNN	rrrrrsuruu	APCAAPPPAA	30.813
		...		
1385206479	DDDNDDDNNN	rrrurrrsuu	APAACAPPPA	35.052
0463159827	DNDDDNNDDN	rsrrruurru	APPPAAAACP	35.137
1342760598	DDNDNDDNND	rrsrurruur	APPCPPAAAA	35.142
1385476209	DDDNNNDDDN	rrrusurrru	APAAPPPCAA	35.296
0213657984	DDDDDNNNDN	rrrrruuurs	ACAPPAPAAP	35.430

Crossover: This study employed the precedence preservative crossover (PPX) methodology for crossover. In this methodology, in addition to the two strings representing the chromosomes of the parents ($Parent_1$ and $Parent_2$), two additional strings pass on the precedence relationship based on the two parental permutations to two new offspring while making sure that no new precedence relationships are introduced. A vector, representing the number of operations involved in the problem is randomly filled with elements of the set. This vector defines the order in which the operations are successively drawn from $Parent_1$ and $Parent_2$.

The algorithm starts by initializing an empty offspring. The leftmost operation in one of the two parents is selected in accordance with the order of parents given in the vector. After an operation is selected it is deleted in both parents. Finally, the selected operation is appended to the offspring. This step is repeated until both parents are empty and the offspring contains all operations involved.

Mutation: The population is subjected to mutation operation with a given probability. If the probability holds, the mutation operator selects a random number of genes (rnd = 1, … ,9), and exchanges them in such a way that the same precedence relationships are preserved. Otherwise the population remains unchanged and is copied to the next generation. The mutation operator proposed in this paper exchanges components 0 and 1. The rest of the strings remain the same.

Fitness Evaluation: The fitness function is dependent on the increment in disassembly time. There are three factors, which add up to the disassembly time of a component. The first one is basic disassembly time for component j in sequence seq ($dt_{j,seq}$). In this paper $dt_{j,seq}$ values (in seconds) are given as:

j	0	1	2	3	4	5	6	7	8	9
$dt_{j,seq}$	2	3	3	2	3	4	2	1	3	2

The second function ($ct_{j,seq}$) is the penalty (in seconds) for each travel time to disassemble component j in sequence seq, which includes a function of the distance traveled between the (seq-1)th and seq^{th} sequences and the robot arm speed factor (sf):

$$ct_{j,seq} = \sqrt{(x_{j,(seq-1)}-x_{j,seq})^2+(y_{j,(seq-1)}-y_{j,seq})^2+(z_{j,(seq-1)}-z_{j,seq})^2} \Big/ sf$$

The last criterion in fitness function is the penalty for disassembly method change ($mt_{j,seq}$). For each disassembly method change, the sequence is penalized by 1 second:

$$mt_{j,seq} = \begin{cases} 0, & \text{If no method change is required,} \quad (e.g.\ N \text{ to } N) \\ 1, & \text{If method change is required,} \quad (e.g.\ N \text{ to } D) \end{cases}$$

In addition, the algorithm searches for a “recycling pair” and does not penalize the sequence if the two adjacent components are made of the same material and if they are both demanded for recycling.

Let T_{seq} denote the cumulative disassembly time after the disassembly operation in sequence seq is completed for component j:

$$\begin{aligned} T_{seq} &= T_{seq-1} + dt_{j,seq} + ct_{j,seq} + mt_{j,seq}, \text{ for } seq = 0, \dots, n\text{-}2 \\ T_{seq} &= T_{seq-1} + dt_{j,seq}, \text{ for } seq = n\text{-}1 \end{aligned} \qquad (1)$$

In this proposed GA model, the objective is to minimize the total fitness function (F) by minimizing (i) the traveled distance, (ii) the number of disassembly method changes, and (iii) by combining the identical-material components together, eliminating unnecessary disassembly operations. Let $F(ch,gn)$ denote the total fitness for chromosome ch in generation gn. Hence, total time to disassemble all the components can be calculated as follows:

$$F(ch,gn) = \sum_{seq=0}^{n-1} dt_{j,seq} + \sum_{seq=0}^{n-2} ct_{j,seq} + \sum_{seq=0}^{n-2} mt_{j,seq}, \ \forall\, j, j = 0, \dots, n\text{-}1. \qquad (2)$$

Selection and Regeneration Procedure: After every generation, the chromosomes obtains a certain expectation depending on their fitness values. A roulette wheel is then implemented to select the sequence of parents that will be included in the next generation

(the higher the fitness value the higher the chance to be selected). This method aims at allowing the parents in the current generation to be selected for the next generation without getting trapped in the local optima. In addition, a new population is generated eliminating the weak chromosomes.
Termination: The execution of GA terminates if the number of generations reaches up to a maximum value (100 in our example).

3.3 Genetic Algorithm Model Assumptions

Proposed model assumes that the end effector speed for the robot arm is a constant value of 25 cm/sec. In addition, the time spent for robot arm angle change (for all three angles) is assumed to be embedded in the disassembly time for each component. In addition, every component is assumed to have one joint that connects the component with the rest of the product structure.

4. NUMERICAL EXAMPLE

Consider the product in Figure-1. The crossover and mutation probabilities are assumed to be 0.60 and 0.005 respectively. Initial population provided in Table 1 consists of 100 chromosomes (*ncr* = 100). After the GA is run, only one optimal solution is obtained in the final population with a fitness function value of 26.0248 seconds.

Table-2. Final population for the numerical example

Sequence	***F(ch,gn)***
1 3 2 8 0 6 9 5 4 7 DDDDDDNNNN rrrrrruusu APCAAPAAPP	26.0248 secs

The time to calculate the optimal solution (Table-2) took 2. 4180 seconds using the exhaustive search algorithm where as this time was increased to 2.5576 seconds for the genetic algorithm. The solution was reached at the 6^{th} generation. The initial solution of 100 chromosomes included the four duplicates, identical chromosomes, where as the rest was unique sequences. It is important to note that the sequence is found in a few iterations even though it did not take place in the initial random population.

As for the average values of 100 runs, the average time to obtain an optimal solution took 2.5576 seconds, and the optimal solution was reached at the 16.56^{th} generation at the average. The code is written in Matlab (version 7.7.0.471(R2008b)) using Genetic Algorithm and Direct Search Toolbox. The code was run on a computer with a Processor Intel Core 2 Duo CPU P8400 2.26 GHz, 4.00 GB RAM (Table-4).

5. CONCLUSIONS

In this paper, a genetic algorithm model was presented in order to determine the optimal disassembly sequence of a given product. The model provides quick and reliable input to the disassembly scheduling environments. As also stated by Keung *et al*. [2], GAs do not make unrealistic assumptions such as linearity, convexity and/or differentiability. This adds further importance to the proposed model and makes it even more desirable. For the example considered, the algorithm provided optimal disassembly sequence in a short execution time. The algorithm is practical, as it is easy to use, considers the precedence relationships and additional constraints in the product structure and is easily applicable to problems with multiple objectives.

References

[1] Valenzuela-Rendon, M. and Uresti-Charre, E. *A Non-generational Genetic Algorithm for Multiobjective Optimization*. in *Seventh International Conference on Genetic Algorithms*, 1997.

[2] Keung, K.W., Ip, W.H., and Lee, T.C. *The Solution of a Multi-Objective Tool Selection Model Using the GA Approach. International Journal of Advanced Manufacturing Technology* 18: p. 771-777, 2001.

[3] Lazzerini, B. and Marcelloni, F. *A Genetic Algorithm for Generating Optimal Assembly Plans. Artificial Intelligence in Engineering* 14: p. 319-329, 2000.

[4] Kongar, E. and Gupta, S.M. *Disassembly Sequencing using Genetic Algorithm. International Journal of Advanced Manufacturing Technology* 30(5-6): p. 497-506, 2006.

[5] McGovern, S.M. and Gupta, S.M. *A Balancing Method and Genetic Algorithm for Disassembly Line Balancing. European Journal of Operational Research* 179(3): p. 692-708, 2007.

[6] Ilgin, M.A. and Gupta, S.M. *Environmentally Conscious Manufacturing and Product Recovery (ECMPRO): A Review of the State of the Art. Journal of Environmental Management* 91(3): p. 563-591, 2010.

[7] Gungor, A. and Gupta, S.M. *Issues in Environmentally Conscious Manufacturing and Product Recovery: A Survey. Computers and Industrial Engineering* 36(4): p. 811-853, 1999.

[8] Lee, D.-H., Kang, J.-G., and Xirouchakis, P. *Disassembly Planning and Scheduling: Review and Further Research. Journal of Engineering Manufacture* 215(B5): p. 695-709, 2001.

[9] Lambert, A.J.D. and Gupta, S.M. *Disassembly Modeling for Assembly, Maintenance, Reuse, and Recycling*. 2005, Boca Raton, Florida: CRC Press.

[10] Sanderson, A.C., Homem de Mello, L.S., and Zhang, H. *Assembly Sequence Planning. Assembly Sequence Planning." AI Magazine* (Spring 1990): p. 62-81, 1990.

[11] Seo, K.-K., Park, J.-H., and Jang, D.-S. *Optimal Disassembly Sequence Using Genetic Algorithms Considering Economic and Environmental Aspects. International Journal of Advanced Manufacturing Technology* 18: p. 371-380, 2001.

[12] Bierwirth, C., Mattfeld, D.C., and Kopfer, H., *On Permutation Representations for Scheduling Problems*. in *Parallel Problem Solving from Nature -- PPSN IV, Lecture Notes in Computer Science*, Voigt, H.M., et al., Editors, 1996, Springer-Verlag: Berlin, Germany, p. 310-318.

[13] Bierwirth, C. and Mattfeld, D.C. *Production Scheduling and Rescheduling with Genetic Algorithms. Evolutionary Computation* 7(1): p. 1-18, 1999.

[14] Shimizu, Y., Tsuji, K., and Nomura, M. *Optimal Disassembly Sequence Generation Using a Genetic Programming. International Journal of Production Research* 45(18-19): p. 4537–4554, 2007.

[15] Hui, W., Dong, X., and Guanghong, D. *A Genetic Algorithm For Product Disassembly Sequence Planning. Neurocomputing* 71: p. 2720– 2726, 2008.

[16] Torres, F., Gil, P., Puente, S.T., Pomares, J., and Aracil, R. *Automatic PC Disassembly for Component Recovery. International Journal of Advanced Manufacturing Technology* 23(1-2): p. 39-46, 2004.

[17] Torres, F., Puente, S.T., and Aracil, R. *Disassembly Planning Based On Precedence Relations Among Assemblies. International Journal of Advanced Manufacturing Technology* 21(5): p. 317-327, 2003.

[18] Pomares, J., Puente, S.T., Torres, F., Candelas, F.A., and Gil, P. *Virtual Disassembly Of Products Based On Geometric Models. Computers in Industry* 55(1): p. 1-14, 2004.

[19] Gil, P., Pomares, J., Puente, S.V., Diaz, C., Candelas, F., and Torres, F. *Flexible Multi-Sensorial System For Automatic Disassembly Using Cooperative Robots. International Journal of Computer Integrated Manufacturing* 20(8): p. 757-772, 2007.

[20] Torres, F., Puente, S., and Diaz, C. *Automatic Cooperative Disassembly Robotic System: Task Planner To Distribute Tasks Among Robots. Control Engineering Practice* 17(1): p. 112-121, 2009.

AN EVALUATION OF THE EFFECT OF SENSOR EMBEDDED PRODUCTS IN PRODUCT DISASSEMBLY WITH PRECEDENCE RELATIONSHIPS

Mehmet Ali ILGIN, Northeastern University, Boston, MA 02115, 617-373-7635, ilgin.m@neu.edu
Surendra M. GUPTA, Northeastern University, Boston, MA 02115, 617-373-4846, gupta@neu.edu

ABSTRACT

A promising solution approach to deal with the uncertainty associated with disassembly yield is sensor embedded products (SEPs) which involve sensors implanted during the production process. This study presents a quantitative assessment of the impact of SEPs on the various performance measures of a kanban-controlled air conditioner disassembly line. First, separate design of experiments studies based on orthogonal arrays are carried out for the cases with and without SEPs. Then, the results of paired-t tests comparing two cases based on different performance measures are presented.

1. INTRODUCTION

Disassembly is an important process in product recovery since it allows for the selective separation of desired parts and materials. Due to its high productivity and suitability for automation, disassembly line is the most suitable layout for disassembly operations. A disassembly line inherits more uncertainty compared to an assembly line. The quality, quantity and arrival time of components used in the stations of an assembly line can be controlled by imposing strict conditions on component suppliers. However, there are no such strict conditions on the quality, quantity and arrival time of end-of-life (EOL) products which are dismantled into components through a disassembly line. This increases the uncertainty associated with the condition of the components to be disassembled from EOL products (Ilgin and Gupta [2]).

The use of sensor embedded products (SEPs) is a promising approach to deal with the uncertainty in disassembly yield. SEPs involve sensors implanted during the production process. By monitoring critical components of a product, these sensors facilitate the data collection process. The data collected through sensors can be used to predict the component or product failures during product lives while allowing for accurate estimation of remaining lives and conditions of components at the EOL phase. Moreover, the information provided by sensors regarding any defective, replaced or missing component prior to disassembly of an EOL product provides important savings in testing, disassembly, disposal, backorder and holding costs.

There is a vast amount of literature on the use of sensor-based technologies on after-sale product condition monitoring. Starting with the study of Scheidt and Shuqiang [7], different methods of data acquisition from products during product usage were presented by the researchers (Karlsson [3]; Klausner, et al. [5]; Petriu, et al. [6]). In all of these studies, the main idea is the use of devices with memory to save monitoring data generated during the product usage. Although most of these studies focus on the development of SEP models, only few researchers presented a cost-benefit analysis. Klausner, et al. [4] analyzed the trade-off between the higher initial manufacturing cost caused by the use of an electronic data log in products and cost savings from the reuse of used motors. Simon, et al. [8] improved the cost-benefit analysis of Klausner, et al. [4] by considering the limited life of a product design. They showed that, in that case, servicing provides more reusable components compared to EOL recovery of parts. Vadde, et al. [10] investigated the effectiveness of embedding sensors in computers by comparing several performance measures in the two scenarios-with embedded sensors and without embedded sensors. The performance measures considered include average life cycle cost, average maintenance cost, average disassembly cost, and average downtime of a computer. However, they do not provide a quantitative assessment of the impact of SEPs on these performance measures. Moreover, since only one component of a computer (hard disk) was considered, the disassembly setting does not represent the complexity of a disassembly line which is generally used to disassemble EOL computers. By extending Vadde, et al. [10], Ilgin and Gupta [1] analyzed the effect of SEPs on the performance of an EOL computer disassembly line which is used to disassemble three components from EOL computers, namely, memory, hard disk and motherboard. Due to relatively simple structure of an EOL computer, they did not consider the precedence relationships among the components. However, disassembly of a particular component is restricted by one or more components in some products. That is why, these products are disassembled according to a route determined based on the precedence relationships.

In this study, we extend Ilgin and Gupta [1] by investigating the quantitative impact of SEPs in case of product disassembly with precedence relationships. Specifically, we consider a kanban controlled air conditioner (AC) disassembly line since an AC involves several precedence relationships among its components. Separate design of experiments (DOE) studies based on orthogonal arrays are performed for the cases with and without SEPs. In the calculation of various performance measure values under different experimental conditions, detailed discrete event simulation models of both cases are used. Then, the results of paired-t tests comparing two cases based on different performance measures are presented.

2. SYSTEM DESCRIPTION

We consider a six-station EOL AC disassembly line. The components disassembled at different stations and the precedence relationships among the components can be seen in Table 1. Although an AC consists of nine components, there is a customer demand for only four components, namely, control box, motor, fan and compressor. Disassembly times at stations, demand inter-arrival times for components and EOL AC inter-arrival times are all exponentially distributed. A conventional AC visits all stations. Following the disassembly at each station, components are tested. The testing times are normally distributed with the means and standard deviations presented in Table 2. A sensor embedded AC visits only the stations which are responsible for the disassembly of functional components and predecessor components of these components. Moreover, following the disassembly of a component, there is no testing operation due to sensor information on the condition of the component.

A small truck with a load volume of 425 cubic feet is used to dispose of excess product, subassembly and components. Whenever the total volume of the excess product, subassembly and component inventories become equal to the truck volume, the truck loaded with excess inventory is sent to a recycling facility. Any product, subassembly or component inventory which is greater than *maximum inventory level* is assumed to be excess. Component volumes are given in Table 2. The volume of an EOL AC is taken as 3 cubic feet. Multi kanban system (MKS) developed by Udomsawat and Gupta [9] is used to control the disassembly line.

Table 1. Precedence relationships among the parts of an AC.

Part Name	Code	Precedence Relationship	Station
Evaporator	A	-	1
Control Box	B	-	2
Blower	C	A,B	3
Air Guide	D	A,B,C	
Motor	E	A,B,C,D	4
Condenser	F	-	5
Fan	G	F	
Protector	H	-	6
Compressor	I	H	

Table 2. Part characteristics.

Part Name	Code	Testing Time		Volume(cft)	Demanded?	Disposal Classification
		Mean	Std. Dev.			
Evaporator	A	-	-	0.181	No	Copper scrap
Control Box	B	5	1	0.083	Yes	Waste
Blower	C	-	-	0.081	No	Waste
Air Guide	D	-	-	0.406	No	Waste
Motor	E	10	2	0.114	Yes	Waste
Condenser	F	-	-	0.327	No	Copper scrap
Fan	G	2	0.2	0.271	Yes	Waste
Protector	H	-	-	0.012	No	Waste
Compressor	I	10	2	0.114	Yes	Waste

3. DESIGN OF EXPERIMENTS STUDY

In this section, we present a comprehensive and quantitative evaluation of the SEPs on the performance of the AC disassembly line by comparing the case with SEPs against the case without SEPs under different experimental conditions. Table 3 presents the factors and factor levels. A full factorial design with 38 factors requires an extensive number of experiments. Therefore, experiments were designed using orthogonal arrays (OAs) which allow for the determination of main effects by making a minimum number of experiments. Specifically, L81 OA was chosen since it requires 81 experiments while accommodating 40 factors with three levels. Discrete event simulation (DES) models of both cases were developed using Arena 11 to determine profit value together with various cost and revenue parameters for each experiment. Each DES experiment was carried out for 60480 minutes, the equivalent of six months with one eight hour shift per day. The following formula is used in the DES models for the calculation of profit values.

$$Profit = \underbrace{(SR+CR+SCR)}_{\text{Total Revenue}} - \underbrace{(HC+BC+DC+DPC+TC+TPC)}_{\text{Total Cost}} \quad (1)$$

The different cost and revenue components used in Equation (1) can be defined as follows. SR is the total revenue generated by the component sales during the simulated time period (STP). CR is the total revenue generated by the collection of EOL ACs during the STP. SCR is the total revenue generated by selling scrap components during the STP. HC is the total holding cost of components, EOL ACs and subassemblies during the STP. BC is the total backorder cost of components during the STP. DC is the total disassembly cost during the STP. DPC is the total disposal cost of components, EOL ACs and subassemblies during the STP. TC is the total testing cost during the STP. TPC is the total transportation cost during the STP.

In each AC, evaporator and condenser are sold as copper scrap whereas chassis and metal cover are sold as steel scrap. All the other components are regarded as waste components. In order to determine total weight of small components such as screws, cables, total weight of the main components of an AC is multiplied with a *small component weight factor*. In order to calculate the disposal cost of a waste component, the weight is multiplied with the *unit disposal cost*. Disposal cost for subassemblies and products are calculated by multiplying the total weight of waste components in subassembly or product with the *unit disposal cost*. Disposal cost for subassemblies and products are increased by a factor called *disposal cost increase factor for EOL ACs*. Scrap revenue for evaporator and condenser is calculated by multiplying the weight with the *unit copper scrap revenue*. In the calculation of scrap revenue for subassemblies and products, total weight of evaporator and condenser is multiplied with the *unit copper scrap revenue* whereas the total weight of chassis and metal cover is multiplied with the *unit steel scrap revenue*. Scrap revenue for subassemblies and products are decreased by a factor called *scrap revenue decrease factor for EOL ACs*. The total weight of chassis and metal cover is taken as 6 lbs. The time required to retrieve information from the sensors is assumed to be testing time. Duration of this retrieval process is taken as 20 seconds per AC. In the calculation of transportation cost, the operating cost associated with each trip of the truck is assumed to be $50. For each EOL AC, the facility demands a $20 collection fee.

4. RESULTS

Based on the results of DOE studies, various paired-t tests were performed on different performance measures. Table 4 presents the 95% confidence interval, t-value and p-value for each test. According to this table, SEPs achieve statistically significant savings in holding, backorder, disassembly, disposal, testing and transportation costs. Moreover, there are statistically significant improvements in total revenue and profit for the case of SEPs.

Table 3. Factor levels.

No.	Factor	Unit	Levels		
			1	**2**	**3**
1	Disposal cost increase factor for EOL ACs	%	5	10	15
2	Scrap revenue decrease factor for EOL ACs	%	5	10	15
3	Mean demand rate for Control Box	parts/hour	10	15	20
4	Mean demand rate for Motor	parts/hour	10	15	20
5	Mean demand rate for Fan	parts/hour	10	15	20
6	Mean demand rate for Compressor	parts/hour	10	15	20
7	Mean arrival rate of EOL ACs	products/hour	10	20	30
8	Mean disassembly time for station 1	minutes	0.75	1	1.25
9	Mean disassembly time for station 2	minutes	0.75	1	1.25
10	Mean disassembly time for station 3	minutes	0.75	1	1.25
11	Mean disassembly time for station 4	minutes	0.50	0.75	1
12	Mean disassembly time for station 5	minutes	0.75	1	1.25
13	Mean disassembly time for station 6	minutes	1	1.5	2
14	Backorder cost rate	%	40	60	80
15	Disassembly cost per minute	$	1	2	3
16	Testing cost per minute	$	0.40	0.50	0.60
17	Holding cost rate	%	10	20	30
18	Weight for Evaporator	lbs.	8	12	16
19	Weight for Motor	lbs.	6	10	14
20	Weight for Condenser	lbs.	12	15	18
21	Weight for Compressor	lbs.	6	10	14
22	Price for Control Box	$	30	60	90
23	Price for Motor	$	50	100	150
24	Price for Fan	$	15	30	45
25	Price for Compressor	$	50	75	100
26	Unit disposal cost	$/lbs	0.30	0.40	0.50
27	Unit copper scrap revenue	$/lbs	0.40	0.60	0.80
28	Unit steel scrap revenue	$/lbs	0.15	0.20	0.25
29	Maximum inventory level	parts, subassemblies, products	5	10	15
30	Small component weight factor	%	5	10	15
31	Probability of a non-functional Control Box	%	10	20	30
32	Probability of a non-functional Motor	%	10	20	30
33	Probability of a non-functional Fan	%	10	20	30
34	Probability of a non-functional Compressor	%	10	20	30
35	Probability of a missing Control Box	%	5	10	15
36	Probability of a missing Motor	%	5	10	15
37	Probability of a missing Fan	%	5	10	15
38	Probability of a missing Compressor	%	5	10	15

5. CONCLUSIONS

Sensors embedded in products can deal with the uncertainty associated with the disassembly yield by providing information on the condition of the components prior to disassembly. In this study, we analyzed the impact of SEPs on the various performance measures of an AC disassembly line. First, two separate DOE studies based on OAs were carried out for the cases with and without

SEPs. Then paired-t tests were performed in order to compare two cases for various performance measures. According to test results, SEPs achieve significant reductions in holding, backorder, disassembly, disposal, testing and transportation costs while increasing total revenue and profit.

Table 4. Paired-t test results for various performance measures.

Performance Measure	95% Confidence Interval on Mean Difference (Sensor –No Sensor)	t-value	p-value
Holding Cost	(-410.850, -95.004)	-3.19	0.002
Backorder Cost	(-64751.7, -47287.2)	-12.77	0.000
Disassembly Cost	(-46141.9, -32744.6)	-11.72	0.000
Disposal Cost	(-34185.6, -22951.7)	-10.12	0.000
Test Cost	(-175859, -150695)	-25.83	0.000
Transportation Cost	(-2223.76, -1708.21)	-15.18	0.000
Total Cost	(-313978, -265077)	-23.56	0.000
Total Revenue	(328616, 448101)	12.94	0.000
Profit	(601697, 754075)	17.71	0.000

REFERENCES

[1] Ilgin, M. A. and Gupta, S. M., "Analysis of a Kanban Controlled Disassembly Line with Sensor Embedded Products," in *Proceedings of the 2009 Northeast Decision Sciences Institute Conference*, Uncasville, Connecticut, 2009, 555-560.

[2] Ilgin, M. A. and Gupta, S. M., "Environmentally conscious manufacturing and product recovery (ECMPRO): A review of the state of the art," *Journal of Environmental Management,* 2010, 91, 563-591.

[3] Karlsson, B., "A distributed data processing system for industrial recycling," in *Proceedings of the 1997 IEEE Instrumentation and Measurement Technology Conference*, Ottowa, Canada, 1997, 197-200

[4] Klausner, M., Grimm, W. M., and Hendrickson, C., "Reuse of electric motors in consumer products," *Journal of Industrial Ecology,* 1998, 2, 89-102.

[5] Klausner, M., Grimm, W. M., and Horvath, A., "Integrating product takeback and technical service," in *Proceedings of the 1999 IEEE International Symposium on Electronics and the Environment*, Danvers, MA, 1999, 48-53.

[6] Petriu, E. M., Georganas, N. D., Petriu, D. C., Makrakis, D., and Groza, V. Z., "Sensor-based information appliances," *IEEE Instrumentation & Measurement Magazine,* 2000, 3, 31-35.

[7] Scheidt, L. and Shuqiang, Z., "An approach to achieve reusability of electronic modules," in *Proceedings of the 1994 IEEE International Symposium on Electronics and the Environment*, San Francisco, CA, 1994, 331-336.

[8] Simon, M., Bee, G., Moore, P., Pu, J.-S., and Xie, C., "Modelling of the life cycle of products with data acquisition features," *Computers in Industry,* 2001, 45, 111-122.

[9] Udomsawat, G. and Gupta, S. M., "Multikanban system for disassembly line," in *Environment Conscious Manufacturing*, S. M. Gupta and A. J. D. Lambert, Eds. Boca Raton, FL: CRC Press, 2008, 311-330.

[10]Vadde, S., Kamarthi, S., Gupta, S. M., and Zeid, I., "Product Life Cycle Monitoring via Embedded Sensors," in *Environment Conscious Manufacturing*, S. M. Gupta and A. J. D. Lambert, Eds. Boca Raton, FL: CRC Press, 2008, 91-103.

PREEMPTIVE GOAL PROGRAMMING FOR SOLVING THE MULTI-CRITERIA DISASSEMBLY-TO-ORDER PROBLEM UNDER STOCHASTIC YIELDS, LIMITED SUPPLY, AND QUANTITY DISCOUNT

Amre Z. Massoud, Northeastern University, Boston, MA, 02115, (617) 373-7635, massoud.a@neu.edu
Surendra M. Gupta*, Northeastern University, Boston, MA, 02115 (617) 373-4846, gupta@neu.edu
(*corresponding author)

ABSTRACT

This paper considers the disassembly-to-order (DTO) problem where end-of-life (EOL) products are purchased for disassembly in order to satisfy the demand for specified numbers of components. Previous work in the literature solved the DTO problem with multiple uncertainties under a single objective while others solved the multi-criteria model with a single uncertainty. This paper develop a multi-criteria DTO model that would takes into consideration multiple system uncertainties and variability. The multi-criteria DTO model was solved using preemptive goal programming (PGP). The main objective was to determine optimal number of take-back EOL products in every period from each supplier in order to satisfy the demand of components and materials while trying to achieve the aspiration levels of multiple goals. A numerical example is considered to illustrate the model approach.

INTRODUCTION

In recent years, the idea of properly managing waste generated from end-of-life (EOL) products has caught the attention of government officials and law makers. Throughout the world, waste generated from EOL electronic products has grown incredibly large over the past few years. Fully functional electronic products such as personal computers, televisions, batteries, appliances, and cell phones are being dumped into landfills every year. Recent data and studies from the US Environmental Protection Agency (EPA) show an increase in the amount of waste generated by residents, businesses, and institutions. In 2006, the amount of Municipal Solid Waste (MSW) generated in the US was more than 251 million tons which translates into 4.6 pounds of waste per person per day.

Technology plays an essential role in people's lives because it has simplified their lives in many ways. Additionally, the cost of technology has been declining steadily over the years making them more readily accessible to wider range of consumers. As a result, demand in recent years for products with higher technologies has surged. This has motivated manufacturing companies to continuously develop newer products with technologies that would surpass pervious products. Consequently, consumers are tempted to upgrade their old products and acquire new products even though the older products are still in good functioning condition. This rush to upgrade the products with lower technologies has shortened the products lives and has led to the premature disposal of those products.

In order to minimize the negative impact on the environment, it is vital to properly manage the waste generated from the premature disposal of EOL products. Therefore, new laws and regulations have been established by governments that aim to control the amount of waste being sent to landfills. In addition, consumers' awareness has forced manufacturers to become more environmentally conscious. As a result, manufacturers have thought of different ways to deal with waste generated from EOL products. Some of the options include: remanufacturing, reusing, recycling, or disposing of EOL products. If EOL products are remanufactured, reused, or recycled; then manufacturers can help decrease the rate of depletion of virgin nature resources and reduce their dependency on the natural resources in addition to cutting on the amount of harmful waste sent to landfills. However, disposal of EOL products contributes to the problem and therefore is used as a last resort.

Manufacturers first need to estimate the number of EOL products required to fulfill the demand of components. After that, they need to obtain the EOL products through the disassembly process. Then, the EOL products need to be disassembled into individual components and subassemblies. However, a lot of uncertainties exists in the process which complicates the process and make it difficult to know in advance the exact number of EOL products needed for disassembly. Models and techniques have been developed to address some of the uncertainties in the process however, they suffer from some limitation. Most of the previous models consider only one uncertainty which was the yield of retuned products. However, there are a number of other uncertainties that exist in the process such as the number of suppliers offering EOL products, ready-to-use components offered by an independent outside supplier, inventory capacity limit, and quantity discounts offered by suppliers. A limited number of models considered more than one uncertainty but solved the model under a single objective function.

Our paper focuses on the disassembly-to-order (DTO) system where a variety of EOL products are purchased for disassembly. We develop a multi-criteria DTO model that takes into consideration the above mentioned uncertainties and solve it using preemptive goal programming (PGP). The main objective is to determine the best combination of take-back EOL products to be purchased from every supplier for disassembly in order to satisfy the demand of components and materials while trying to achieve the aspiration levels of multiple goals. A numerical example is considered to illustrate the model approach.

LITERATURE REVIEW

The area of disassembly has sparked a lot of interest recently. Some researchers have examined the disassembly process and how it affects the environment while others focused on the financial aspects of the process. Gungor and Gupta [1] presented a comprehensive study of the different issues in environmentally conscious manufacturing and product recovery. Imtanavanich and Gupta [2] considered the DTO problem where the main objective was to determine the optimal number of returned products to be disassembled in order to fulfill the demand for a specified number of parts. Massoud and Gupta [3] considered the multi-period DTO problem under stochastic yields, limited supply, and quantity discount. The main objective was to determine the best combination of EOL products to be taken back in order to maximize total profit. Inderfurth and Langella [4] developed two heuristics for solving DTO problems with stochastic yields. The heuristic procedure transformed the stochastic yields into their equivalent deterministic yields. Imtanavanich and Gupta [5] developed a multi-criteria decision making approach for a DTO system under product deterioration and stochastic yields. Imtanavanich and Gupta [6] used a multi-criteria decision making approach in order to model the DTO problem with stochastic yields and used heuristic procedures to convert the stochastic yields into their deterministic equivalents. Kongar and Gupta [7] presented a multi-criteria decision making approach for DTO system. Goal programming was used to determine the best combination of EOL products to be disassembled in order to fulfill the demand for items and materials under a number of financial, environmental, and physical constraints.

PROBLEM DEFINITION

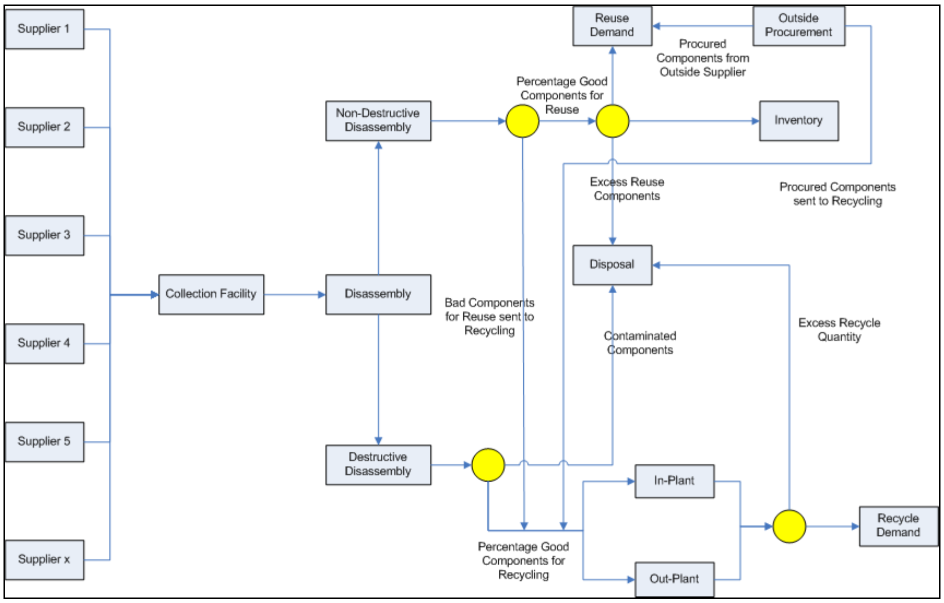


Figure 1: Disassembly-to-Order System

In a DTO problem, a variety of returned EOL products are considered for disassembly in order to satisfy the demand for certain components. The DTO process starts with the purchased of EOL products from a predetermined number of suppliers (see figure 1). The EOL products are then sent to the collection facility to be inspected and prepared for disassembly. Once they have been cleaned and inspected, they are sent to the disassembly process where components are separated using one of two processes: non-destructive or destructive disassembly process. Reuse demand along with components sent to storage are going to be satisfied with components from the non-destructive disassembly process. The number of stored components is limited by the available storage space. Recycling demand is going to be satisfied with components from the destructive disassembly process.

The yield of components from non-destructive disassembly is stochastic. Components obtained from non-destructive disassembly are sorted and tested for functionality. Next, they are divided into two types according to their functionality: good non-destructive components and bad non-destructive components. Again, the percentage of functional components is stochastic. Only good non-destructive components are used to satisfy the demand for reused components and storage. Since demand shortage is not allowed, all demand must be met. If reuse demand cannot be met from the good non-destructive components, additional components may be purchased at a premium cost from an outside supplier. Damaged components along with components that are not worthy of non-destructive disassembly are considered bad non-destructive components. Components obtained from destructive disassembly are also inspected and divided into two types: good destructive components and bad (contaminated) destructive components. Good destructive components along with bad non-destructive components are sent to the recycling process. Components are recycled either in-plant or out-plant in order to satisfy the recycling demand. In-plant recycling process has a capacity limit on the number of components that can be processed. Components are going to be sent first to the in-plant recycling process since it is cheaper than out-plant recycling. If the demand is greater than the in-plant processing capacity, additional components are sent to the out-plant process. Finally, components that exceed the demand for reuse, recycle, and storage are sent to disposal along with the contaminated components from the recycling process.

In this paper, we develop a multi-criteria DTO model which takes into consideration the different system uncertainties. PGP is used to solve the multi-criteria DTO model. The main objective is to determine the best combination of take-back EOL products to be purchased from every supplier that would satisfy the demand while achieving the aspiration levels of multiple goals.

NUMERICAL EXAMPLE

This section presents a numerical example to illustrate the application of the multi-criteria DTO model. The main objective is to determine the best combination of take-back EOL products purchased from every supplier that would satisfy the demand while achieving the aspiration levels of 4 different goals. In this example, we have 2 suppliers, 6 products, and 10 components. Each supplier is going to be offering 3 different types of products and each product will consist of 8 different components. The price and condition of products will differ across supplier and each supplier is offering a different discount schedule that is based on the total dollar amount of products purchased. Input data for our model is summarized in tables 1 though 5.

Table 1: Product Purchase Price & Capacity

Product	Supplier 1		Supplier 2	
	Purchase Price	Capacity	Purchase Price	Capacity
1	$30	250	$39	486
2	$21	352	$30	459
3	$34	381	$32	268

Table 2: Component Yields from Both Suppliers

		A	B	C	D	E	F	G	H	I	J
Supplier 1	**Produc 1**	0.00	0.67	0.61	0.76	0.79	0.53	0.00	0.52	0.53	0.65
	Produc 2	0.66	0.00	0.58	0.74	0.76	0.58	0.64	0.27	0.00	0.29
	Produc 3	0.75	0.72	0.51	0.00	0.00	0.52	0.59	0.53	0.50	0.62
Supplier 2	**Produc 1**	0.00	0.72	0.63	0.81	0.84	0.54	0.00	0.53	0.57	0.67
	Produc 2	0.69	0.00	0.66	0.87	0.90	0.65	0.66	0.31	0.00	0.33
	Produc 3	0.78	0.85	0.59	0.00	0.00	0.60	0.61	0.59	0.60	0.70

Table 3: Discount Schedule Supplier 1

Discount Schedules	Discount Rate
0 < 5,000	0.0%
5,000 < 10,000	3.0%
10,000 < 15,000	5.0%
>= 15,000	8.0%

Table 4: Discount Schedule Supplier 2

Discount Schedules	Discount Rate
0 < 5,000	0.0%
5,000 < 10,000	2.5%
10,000 < 15,000	6.0%
>= 15,000	7.5%

Table 5: Component Input Data

Component	Reuse Demand	Reuse Sell	Reuse Perc.	Recycle Demand	Recycle Sell	Recycle Perc.	Holding Cost	Proc. Cost	Non-Dest. Cost	Dest. Cost	Disposal Cost
A	400	$23	80%	280	$13	98%	$3.45	$10	$0.55	$0.35	$2.60
B	400	$23	85%	221	$16	95%	$3.40	$9	$0.50	$0.30	$2.40
C	550	$22	85%	240	$16	99%	$3.35	$8	$0.45	$0.25	$2.70
D	300	$18	80%	220	$13	96%	$3.35	$7	$0.50	$0.30	$2.70
E	310	$22	75%	280	$15	98%	$3.45	$10	$0.60	$0.40	$2.80
F	560	$21	90%	260	$15	97%	$3.40	$8	$0.65	$0.35	$2.50
G	400	$22	85%	240	$17	97%	$3.55	$11	$0.60	$0.30	$2.90
H	550	$20	85%	360	$15	99%	$3.25	$10	$0.55	$0.35	$2.40
I	320	$21	80%	260	$14	99%	$3.35	$8	$0.50	$0.30	$2.55
J	590	$23	80%	360	$15	99%	$3.40	$9	$0.45	$0.25	$2.65

Additional data needed to solve the model: in-plant recycling cost = $1.00, out-plant recycling cost = $1.80, contaminated disposal cost = $2.00, and inventory capacity = 50 for each component.

Lingo 11.0 was used to solve the multi-criteria PGP problem. The total purchase and discount rates for supplier 1 and 2 were $16,170 with 8% discount and $8,825 with 2.5% discount respectively. The quantity of procured components for reuse and recycle, quantity of stored components, and quantity of disposed components from reuse, recycle, and contamination are summarized in table 6. Finally, table 7 presents the aspiration levels for each goal along with the model results at each stage.

Table 6: Model Output for PGP problem

Component	Outside Procurement		Ending Inventory	Disposal Quantity		
	Reuse Procurement	Recycle Procurement		Excess Reuse	Excess Recycle	Contaminated Quantity
A	28	12	0	0	0	4
B	14	15	0	0	0	7
C	108	0	0	0	203	4
D	0	13	21	0	0	5
E	0	59	0	0	0	2
F	107	0	0	0	164	12
G	60	0	0	0	57	7
H	202	0	0	0	175	5
I	53	0	0	0	69	3
J	203	0	0	0	137	4

Table 7: Aspiration Levels for each Goal in PGP Model

	Aspiration Levels	Initial Run	1st Goal	2nd Goal	3rd Goal	4th Goal
Profit	**>= 93,000**	93,523.85	**93,000.00**	**93,000.00**	**93,000.00**	**93,000.00**
Procurement Cost	**<= 8,000**	11,502.22	8,515.47	**8,000.00**	**8,000.00**	**8,000.00**
Purchase Cost	**<= 25,000**	22,014.00	24,347.00	24,558.00	**24,995.00**	**24,995.00**
Disposal Cost	**<= 2,200**	1,651.62	3,714.35	2,309.79	2,185.69	**2,185.69**
Disassembly Cost		2,913.29	3,164.54	3,225.01	3,183.67	3,183.67
Inventory Cost		70.35	60.30	233.55	70.35	70.35
Recycling Cost		4,778.68	5,217.61	5,226.01	5,145.51	5,145.51
Non-Destructive Cost		2,126.52	2,308.41	2,363.68	2,341.98	2,341.98
Destructive Cost		786.77	856.13	861.32	841.69	841.69

CONCLUSION

In this paper, we evaluated the multi-criteria disassembly-to-order (DTO) problem where a variety of returned end-of-life (EOL) products were purchased for disassembly. The main objective was to determine the best combination of EOL products to be purchased from every supplier that would satisfy the demand while achieving the aspiration levels of multiple goals. A number of factors were considered including the condition of returned products, the variety of products from different suppliers, and quantity discounts offered by suppliers. Preemptive Goal Programming (PGP) was used to solve the multi-criteria DTO problem.

REFERENCES

[1] Gungor, A. and Gupta, S. M., "Issues in environmentally conscious manufacturing and product recovery: a survey," *Computers & Industrial Engineering*, Vol. 36, No. 4, 811-853, 1999.

[2] Imtanavanich, P. and Gupta, S. M., "Multi-criteria decision making for disassembly-to-order system under stochastic yields," *Proceedings of the SPIE International Conference on Environmentally Conscious Manufacturing IV*, Philadelphia, Pennsylvania, 147-162, 2004.

[3] Massoud, A. Z. and Gupta, S. M., "Solving the multi-period disassembly-to-order system under stochastic yields, limited supply, and quantity discount," *Proceedings of 2008 ASME International Mechanical Engineering Congress and Exposition*, Boston, MA, 2008.

[4] Inderfurth, K. and Langella, I. M., "Heuristics for solving disassemble-to-order problems with stochastic yields," *OR Spectrum*, Vol. 28, No. 1, 73-99, 2006.

[5] Imtanavanich, P. and Gupta, S. M., "Multi-criteria decision making approach in multiple periods for a disassembly-to-order system under product's deterioration and stochastic yields," *Proceedings of the SPIE International Conference on Environmentally Conscious Manufacturing V*, 10-21, 2005.

[6] Imtanavanich, P. and Gupta, S. M., "Calculating disassembly yields in a multi-criteria decision making environment for a disassembly-to-order system," *Proceedings of the 2005 Northeast Decision Sciences Institute Conference*, Philadelphia, Pennsylvania, 2005.

[7] Kongar, E. and Gupta, S. M., "A multi-criteria decision making approach for disassembly-to-order systems," *Journal of Electronics Manufacturing*, Vol. 11, No. 2, 171-183, 2002.

AN OPTIMIZATION FRAMEWORK FOR ADVANCED DISASSEMBLY/REPAIR-TO-ORDER SYSTEMS WITH REMAINING-LIFE ADJUSTMENT

Onder Ondemir, Northeastern University, Boston, MA, 02115, (617) 373-7635, ondemir.o@neu.edu
Surendra Gupta[1], Northeastern University, Boston, MA 02115, (617) 373-4846, gupta@neu.edu

ABSTRACT

Due to environmental awareness and realization of cost savings, disassembly-to-order (DTO) concept has become popular. One of the main obstacles to making optimal DTO decisions is the uncertainty involved in end-of-life products (EOLPs). This uncertainty is due to the lack of information about the condition and the quantity of EOLPs returned. This uncertainty is removed by advanced disassembly/repair-to-order systems utilizing sensors to monitor the products in their life-cycle. Sensor technology enables remaining life estimation, thus allows advanced DTO models to deal with sophisticated component and product demands with remaining life adjustment.

This paper presents an optimization framework for advanced disassembly/repair-to-order (ADRTO) systems. The method is compared with a TABU search based heuristic algorithm.

INTRODUCTION

An ADRTO system can be considered as an extension to traditional disassembly-to-order (DTO) system. DTO is a process, in which end-of life products (EOLPs) are disassembled in order to fulfill the demand for materials and reusable components. DTO systems recently became popular with the increased public awareness on environmental issues such as depletion of landfills, exhaustion of virgin resources, global warming etc. Environmental consciousness triggered the use of recycled materials and reusable components, and created a demand for them. Thus, traditional supply-demand balance problem was reproduced for used components and recycled materials that are recovered from EOLPs. DTO models address this very problem: optimal planning of disassembly operations. In other words, DTO models try to determine the optimal number of EOLPs to be disassembled in order to satisfy the system criteria (minimum cost, maximum profit, etc.). In fact, this process contains a lot of uncertainties because neither quality, nor the quantity of EOLPs is known before disassembly. Many academicians put a lot of efforts to address this limitation of traditional DTO systems'. For more information about DTO, uncertainties involved in DTO and solution approaches [1-11] can be useful.

Life-cycle data bring clarity to the EOL operations and are used to determine the remaining life of the components [12]. These data are captured using embedded sensors and radio frequency identification (RFID) tags. RFID has long been used to gather a history or trace of object movements [13]. RFID technology can be introduced as an enabler of product lifecycle management (PLM) business, by enhancing the traceability of the product throughout its value chain via automatic identification, enabling the collection of product usage information during its life cycle, and facilitating the integration of product lifecycle information and knowledge [14]. Ondemir and Gupta [15] proposed a mathematical DTO model utilizing life-cycle data in order to fulfill remaining life time based sophisticated component demands. In its follow up papers [16, 17], authors extended the model in order to meet the sophisticated product demands by using repair option, and presented economic justification of establishing advanced

[1] **Corresponding author**. Laboratory for Responsible Manufacturing (LRM), Department of MIE, Northeastern University, 360 Huntington Avenue, Boston, MA, 02115 USA

DTO systems in which sensor-embedded EOLPs are disassembled in order to fulfill sophisticated demands.

In this paper, an optimization framework for advanced disassembly/repair-to-order (ADRTO) systems with remaining-life adjustment is proposed.

PROBLEM DEFINITION

ADRTO concept involves disassembly and repair operations in order to fulfill sophisticated component and product demands as well as material demands. It is assumed that ADRTO accepts completely modular EOLPs in which RFID tags and sensors were embedded. By means of remaining-life adjustment, any over qualified components in a repaired product may (or may not) be replaced with a less good, yet satisfying component. In other words, if a product is repaired to have at least two years of remaining-life, any components having more remaining-life may be replaced with the ones having at least two years of life remaining. That way, the over qualified components may be used to fulfill the demands for components with longer remaining-life. This should decrease outside procurement costs since it is assumed that there is a positive correlation between the remaining-life and the procurement cost of a component. On the other hand, this adjustment may incur extra disassembly and assembly cost. Therefore, a replacement (adjustment) is only beneficial when saving obtained outweighs the extra cost.

EOLPs dealt with are completely modular products consisting of a chassis and a number of different types of components based on their models. All components are assembled on the chassis and no interconnection exists among the components. An EOLP may be repaired to make a product of a certain model that is different from its original model. Therefore, some components may be extra (unnecessary). A repair operation involves disassembly of broken, under-qualified (remaining-life deficit) and extra components, and assembly of necessary components that satisfy remaining-life requirements. When an EOLP is to be repaired, repair plan is populated from the data set stored in RFID tags and there is no variability in the repair plan. In other words, repair operations (disassembly and assembly) are not decision variables. On the other hand, remaining-life adjustment is indeed a decision option and cost of this option is dependent on initial repair decision. This dependency is what makes the problem nonlinear.

Nonlinearity makes hard-to-solve pure integer problem even harder-to-solve. For easy computation, a two-stage linear program is developed. This method yields a near optimal solution.

TWO-STAGE OPTIMIZATION FRAMEWORK MODEL

Although generic ADRTO model is a linear deterministic model [16], remaining-life adjustment renders the problem non-linear. Non-linearity occurs in the objective function because of the structure of the EOLPs. Therefore, as a simplified method, a two-stage linear program reducing the problem to two linear sub-problems is developed. Then, as a second approach, a TABU search based heuristic algorithm is developed and the results of these two methods are compared. As conclusions, pros and cons of the two techniques are evaluated and a numerical example is considered to illustrate both approaches.

The goal of the model is to determine which EOLPs to disassemble, repair, or recycle and which components to replace for remaining life adjustment purposes. The model is constructed for completely modular products. In other words, all components are assumed to be independently assembled on a base (chassis). The mathematical model for stage 1 is shown below.

$$\min z_1 = \sum_{i\epsilon I, j\epsilon J}\Big[ind1_i\big(\bar{x}_i c_{1j} a_{ij} + (\bar{x}_i + \bar{y}_i) c_3 f_{ij}\big) + \\ \big(cd_j \sum_{t\epsilon T}\big(ext_{itj} \sum_{m\epsilon M} y_{itm}\big) + ca_j \sum_{t\epsilon T}\big(mis_{itj} \sum_{m\epsilon M} y_{itm}\big) + \big(cd_j + \tag{1}$$

$ca_j) \sum_{t\epsilon T,m\epsilon M} def_{itmj} y_{itm})\Big] + \sum_{i\epsilon I,t\epsilon T,\{j\epsilon J|j\le v\}} \Big[f_{ij} . \sum_{m\epsilon M} ch_{itm} y_{itm} +$
$\sum_{m\epsilon M} ch_{itm} def_{itmj} y_{itm} + ext_{itj} \sum_{m\epsilon M} ch_{itm} y_{itm} +$
$mis_{itj} \sum_{m\epsilon M} \Big(chm_{itm} - (1 - ind1_i) ext_{itj} cd_j +$
$def_{itmj}(cd_j + ca_j)\Big) y_{itm}\Big] + \sum_{j\in J,b\in B} c_{jb} l_{jb} + \sum_{j\in J} cr_j \big(\sum_{i\in I} f_{ij}(\bar{x}_i + \bar{y}_i) +$
$\sum_{b\in B} r_{jb}\big)$

subject to, $\bar{x}_i + \bar{y}_i \le 1,\ \forall i$ (2)

$$\sum_{b\in B} x_{ijb} = \bar{x}_i,\ \forall i,j \quad (3)$$

$$\sum_{t\in T,m\in M} y_{itm} = \bar{y}_i,\ \forall i \quad (4)$$

$$\sum_{i\in I} y_{itm} \ge dp_{tm},\ \forall t,m \quad (5)$$

$$\sum_{\{i\in I|cin_{ijb}=1\}} \big[a_{ij} x_{ijb} + \sum_{t\in T} \big(ext_{itj} \sum_{m\in M} y_{itm}\big) + \sum_{t\in T,m\in M} def_{itmj} y_{itm}\big] - \sum_{i\in I} \big(\sum_{t\in T,m\in M} (rep_{itmjb})\big) + l_{jb} - r_{jb} \ge dc_{jb}, \forall b,j \quad (6)$$

$$\sum_{\{b\in B|b\ge m\}} rep_{itmjb} = y_{itm}\big(mis_{itj} + def_{itmj}\big), \forall i,j,t,m \quad (7)$$

$$\sum_{j\in J} g_{jk}\big(rb_j + \sum_{b\in B} r_{jb}\big) \ge dm_k, \forall k \quad (8)$$

$$rb_j \le \sum_{i\in I} f_{ij}(\bar{x}_i + \bar{y}_i), \forall j \quad (9)$$

Equation (1) defines the objective function by minimizing the sum of total disassembly cost, total recycling cost and total outside component procurement cost, respectively. Equation (2) represents the constraints that assure that EOLP in the inventory can be disassembled, repaired or left untouched. Equation (3) represents constraints that assure complete disassembly. Equation (4) represents constraints that make sure that an EOLP is repaired to produce only one product and that product is evaluated in only one product life-bin. Equation (5) is the set of constraints that ensure that the sophisticated product demand is satisfied by repaired EOLPs. Equation (6) sets a set of constraints to make sure the component demand is met. Constraints defined in Eq. (7) assure all missing and time-deficit components are replaced. Material demand is satisfied by recycling broken and, if necessary, good components. Recycled broken components cannot be more than disassembled broken components. These are defined in Eq. (8) and (9). All variables are non-negative.

The mathematical model for stage 2 is shown below.

$$\min z_2 = z_1 + \sum_{i\in I,j\in J} \Big(\big(cd_j + ca_j\big)(1 - ind2_i) \sum_{b\in B} opt_{ijb}\Big) + \sum_{b\in B,j\in J} c_{jb} \sum_{i\in I} \big(opt_{ijb} - oq_{ijb}\big) \quad (10)$$

subject to, $\sum_{\{b\in B|b<lifebin_{ij}\}} opt_{ijb} \le \bar{y}_i, \forall i,j$ (11)

$$\sum_{b \in B} opt_{ijb} \leq 1, \forall i, j \tag{12}$$

$$\sum_{b \in B} opt_{ijb} \leq soq_{ijm}, \forall i, j \tag{13}$$

$$\sum_{b \in B} oq_{ijb} = \sum_{b \in B} opt_{ijb}, \forall i, j \tag{14}$$

$$\sum_{\{b \in B | b < m\}} oq_{ijb} = 0, \forall i, j \; and \; for \; \{m \in M | \sum_{t \in t} y_{itm} = 1\} \tag{15}$$

Equation (10) defines the objective function stage 2 by minimizing the sum of cost calculated in stage 1, total optional replacement cost and savings obtained by remaining life adjustment, respectively. Equation (11) represents the constraints that assure that remaining life adjustment is performed on repaired EOLPs only. Equation (12) represents constraints that assure that an over qualified component is evaluated in one component life-bin only. Equation (13) represents constraints that make sure that only over qualified components are disassembled. Equation (14) is the set of constraints that ensure that disassembled over qualified components are replaced. Equation (15) sets a set of constraints to make sure assembled components have a remaining life that is suitable with resulting product's life-bin. All variable are non-negative.

TABU SEARCH BASED HEURISTIC ALGORITHM

A TABU search based heuristic algorithm is developed in MATLAB programming language to solve the original non-linear ADRTO problem with remaining life adjustment. The algorithm starts from a basic feasible solution and randomly visits 100 neighboring feasible solutions in each iteration. Termination criteria are two folds. Algorithm stops when it either reaches the 100^{th} iteration or observes 25 non-improving iterations in a row.

NUMERICAL EXAMPLE

To illustrate the methodology, an example involving modular EOLPs is considered. There are 10 different models of a product and 17 components that the model deals with. Each product model is made of a different combination of these components. Original configuration of each product type and the list of components are given in Table 1. Three remaining-life-bins are defined for components. First life-bin holds those components having a remaining-life time of 2 years or less, second life-bin holds those components whose remaining-life are between two and three years. The last bin holds the other components (having 3 years or more remaining life). Same remaining life time ranges are used to define three product life-bins.

TABLE 1: ORIGINAL CONFIGURATION

	Components	*A*			*B*			*C*				*D*		*E*				
		1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**	**4**	**1**	**2**	**1**	**2**	**3**	**4**	**5**
Models	**1**	X			X			X				X		X				
	2		X		X			X				X		X				
	3			X		X		X				X			X			
	4		X		X					X			X				X	
	5	X					X		X					X				
	6	X				X					X	X			X			
	7			X		X				X					X			
	8			X	X					X				X		X		
	9		X				X		X			X					X	
	10			X			X				X		X					X

There are 200 EOLPs (daily return quantity) in the inventory. Non-operable and operable components in EOLPs and remaining life associated with each operable component are recorded into the EOLP database.

Disassembly and procurement costs differ by component groups. Disassembly costs for component groups A, B, C, D, and E are $0, $0.50, $2.00, $1.00, and $1.50, respectively. Procurement costs also depend on the remaining life of the components and are given in Table 2. Recycling costs are 20 Cents for each component in group C and 10 Cents for the others.

TABLE 2: PROCUREMENT COSTS

	Components	A			B			C				D		E				
		1	2	3	1	2	3	1	2	3	4	1	2	1	2	3	4	5
Bins	1	20	40	60	15	15	15	12	18	22	24	5	8	10	30	45	60	75
	2	40	30	70	20	20	20	18	22	24	30	7	9	15	40	55	65	94
	3	50	60	75	25	25	25	24	26	30	32	10	12	25	60	75	80	105

Demands for each type of component and product are shown in Table 3 and Table 4, respectively

TABLE 3: SOPHISTICATED COMPONENT DEMANDS

	Remaining-life-bins		
Components	**Bin 1**	**Bin 2**	**Bin 3**
A1	3	9	9
A2	3	3	6
A3	0	0	15
B1	3	6	12
B2	0	3	6
B3	0	6	6
C1	3	3	12
C2	3	3	15
C3	6	3	6
C4	3	3	12
D1	3	3	15
D2	9	15	15
E1	3	6	9
E2	3	3	9
E3	3	0	15
E4	0	3	15
E5	0	0	21

TABLE 4: SOPHISTICATED PRODUCT DEMANDS

	Remaining-life-bins		
Models	**Bin 1**	**Bin 2**	**Bin 3**
1	1	6	1
2	1	3	4
3	3	3	3
4	1	4	6
5	1	3	12
6	1	3	3
7	1	4	7
8	3	5	3
9	3	3	3
10	2	6	6

RESULTS AND COMPARISON OF THE METHODS

Programs were run on a host computer featuring 1.80 GHz Intel Core2 Duo processor and 3GB memory. Two-stage optimization framework was solved using LINGO 11.0 and obtained a solution costing $6935 in the first stage and remaining life adjustment (stage 2) decreased the cost to $6169. Stage 1 and stage 2 calculations took 262 and 6 seconds, respectively. On the other hand, the TABU search algorithm found a solution with a total cost of $4581.40, but the calculation took 971.32 seconds. Algorithm stopped because maximum number of iterations was reached.

CONCLUSIONS

In this paper, a two-stage optimization framework for advanced disassembly/repair-to-order (ADRTO) systems is presented and compared with a TABU search based heuristic algorithm developed for the same problem. According to the obtained results, TABU search algorithm found a significantly better solution in a longer time. Calculation time is highly dependent on the features of the host computer, programming language and the programming skills of the programmer. Besides, search parameters may affect the calculation time as well.

As a conclusion, the TABU search algorithm developed for this problem was found to be superior to the two-stage optimization framework despite its calculation time drawback.

REFERENCES

1. Gupta, S.M. and K. Taleb, *Scheduling Disassembly.* International Journal of Production Research, 1994. **32**(8): p. 1857-1866.
2. Lambert, A.J.D., *Optimal disassembly of complex products.* International Journal of Production Research, 1997. **35**(9): p. 2509-23.
3. Meacham, A., R. Uzsoy, and U. Venkatadri, *Optimal disassembly configurations for single and multiple products.* Journal of Manufacturing Systems, 1999. **18**(5): p. 311-322.
4. Taleb, K. and S.M. Gupta, *Disassembly of Multiple Product Structures.* Computers & Industrial Engineering, 1997. **32**(4): p. 949-961.
5. Taleb, K., S.M. Gupta, and L. Brennan, *Disassembly of Complex Products with Parts and Materials Commonality.* Production Planning and Control, 1997. **8**(3): p. 255-269.
6. Tang, Y., et al., *Disassembly modeling, planning, and application.* Journal of Manufacturing Systems, 2002. **21**(3): p. 200-217.
7. Gungor, A. and S.M. Gupta, *Disassembly sequence plan generation using a branch-and-bound algorithm.* International Journal of Production Research, 2001. **39**(3): p. 481 - 509.
8. Gungor, A. and S.M. Gupta, *Disassembly line in product recovery.* International Journal of Production Research, 2002. **40**(11): p. 2569 - 2589.
9. Veerakamolmal, P. and S.M. Gupta, *Analysis of design efficiency for the disassembly of modular electronic products.* Journal of Electronics Manufacturing, 1999. **9**(1): p. 79-95.
10. Kongar, E. and S.M. Gupta, *A multi-criteria decision making approach for disassembly-to-order systems.* Journal of Electronics Manufacturing, 2002. **11**(2): p. 171-83.
11. Kongar, E. and S.M. Gupta, *Disassembly to order system under uncertainty.* Omega, 2006. **34**(6): p. 550-561.
12. Rugrungruang, F., *An Integrated Methodology for Assessing Physical & Technological Life of Products for Reuse*, in *Life Cycle Engineering & Management Research Group, School of Mechanical and Manufacturing Engineering*. 2008, The University of New South Wales: New South Wales, Australia. p. 225.
13. Brusey, J. and D.C. McFarlane, *Effective RFID-based object tracking for manufacturing.* International Journal of Computer Integrated Manufacturing, 2009. **22**(7): p. 638 - 647.
14. Cao, H., et al., *RFID in product lifecycle management: a case in the automotive industry.* International Journal of Computer Integrated Manufacturing, 2009. **22**(7): p. 616 - 637.
15. Ondemir, O. and S.M. Gupta. *End-of-Life Decisions Using Product Life Cycle Information*. in *ASME International Mechanical Engineering Congress and Exposition (IMECE 2008)*. 2008. Boston, MA.
16. Ondemir, O., S.M. Gupta, and K. Nakashima. *Disassembly-to-Order Model for Sensor Embedded Products*. in *the Eighth International Conference on EcoBalance*. 2008. Tokyo, Japan.
17. Ondemir, O. and S.M. Gupta, *Cost-Benefit Analysis of Sensor Embedded Products Based Disassembly-to-Order System*, in *Northeast Decision Science Institute 2009*. 2009: Uncasville, CT.

DISASSEMBLY OF SENSOR EMBEDDED PRODUCTS WITH COMPONENT DISCRIMINATING DEMAND

Mehmet Ali ILGIN, Northeastern University, Boston, MA 02115, 617-373-7635, ilgin.m@neu.edu
Surendra M. GUPTA, Northeastern University, Boston, MA 02115, 617-373-4846, gupta@neu.edu

ABSTRACT

There is a high level uncertainty associated with disassembly yield due to existence of unfunctional and/or unneeded components in end of life products. Sensors embedded in critical components of a product can deal with this uncertainty by providing information on the type and condition of the components prior to disassembly. This study presents a quantitative assessment of the impact of sensor embedded products (SEPs) on the various performance measures of a kanban-controlled washing machine disassembly line. First, separate design of experiments studies based on orthogonal arrays are performed for the cases with and without SEPs. Then, the results of paired-t tests comparing two cases based on different performance measures are presented.

1. INTRODUCTION

Increasing environmental awareness and stricter government regulations on the recycling of end-of-life (EOL) products force manufacturers to establish specific facilities for product recovery which can be defined as recovery of materials and parts from returned or EOL products via recycling and remanufacturing. Disassembly is an important process in product recovery since it allows for the selective separation of desired parts and materials (Ilgin and Gupta [2]). Regarding the condition and type of a component in an EOL product, there are several questions which increase the uncertainty in disassembly yield: a) Is the component functional? b) Is the component missing? c) What is the type of the component? If these questions are answered prior to the disassembly of the component, unnecessary disassembly of a non-functional or unneeded component can be avoided. Sensors embedded in products can answer these questions prior to disassembly. By using the sensor information on the condition and type of the components in EOL products, the routing of products through the disassembly line can be changed.

The use of sensor-based technologies on after-sale product condition monitoring is an active research topic. Starting with the study of Scheidt and Shuqiang [8], different methods of data acquisition from products during product usage were presented by the researchers (Karlsson [3, Karlsson [4]; Klausner, et al. [6]; Petriu, et al. [7]). In all of these studies, the main idea is the use of devices with memory to save monitoring data generated during the product usage. Although most of these studies focus on the development of sensor embedded product (SEP) models, only few researchers presented a cost-benefit analysis. Klausner, et al. [5] analyzed the trade-off between the higher initial manufacturing cost caused by the use of an electronic data log in products and cost savings from the reuse of used motors. Simon, et al. [9] improved the cost-benefit analysis of Klausner, et al. [5] by considering the limited life of a product design. They showed that, in that case, servicing provides more reusable components compared to EOL recovery of parts. Vadde, et al. [11] investigated the effectiveness of embedding sensors in computers by comparing several performance measures in the two scenarios-with embedded sensors and without embedded sensors. The performance measures considered include average life cycle cost, average maintenance cost, average disassembly cost, and average downtime of a computer. However, they do not provide a quantitative assessment of the impact of SEPs on these performance measures. Moreover, since only one component of a computer (hard disk) was considered, the disassembly setting does not represent the complexity of a disassembly line which is generally used to disassemble EOL computers. By extending Vadde, et al. [11], Ilgin and Gupta [1] analyzed the effect of SEPs on the performance of an EOL computer disassembly line which is used to disassemble three components from EOL computers, namely, memory, hard disk and mother board. Due to relatively simple structure of an EOL computer, they did not consider the precedence relationships among the components. However, disassembly of a particular component is restricted by one or more components in some products. That is why, these products are disassembled according to a route determined based on the precedence relationships. Another disassembly scenario that was not investigated in Ilgin and Gupta [1] is the component discriminating demand which occurs when customers demand a specific type of a particular component. Due to changes made by customer or service personnel during product usage, it is very difficult to determine the type of a component in a conventional EOL product prior to disassembly. However, the sensors embedded in SEPs can determine the type of components prior to disassembly. This prevents the disassembly of unneeded component types while decreasing the backorder costs of the highly demanded component types.

In this study, we extend Ilgin and Gupta [1] by investigating the quantitative impact of SEPs in case of product disassembly with precedence relationships and component discriminating demand. Specifically, we consider a kanban controlled washing machine (WM) disassembly line. Two separate Design of Experiments (DOE) studies based on Orthogonal Arrays (OAs) are carried out for the cases with and without SEPs. In the calculation of various performance measure values under different experimental conditions, detailed Discrete Event Simulation (DES) models of both cases are used. Then, the results of paired-t tests comparing two cases based on different performance measures are presented.

2. SYSTEM DESCRIPTION

A three-station EOL WM disassembly line is analyzed in this study. The components disassembled at different stations and the precedence relationships among the components can be seen in Table 1. Disassembly times at stations, demand inter-arrival times for components and EOL WM inter-arrival times are all exponentially distributed. There are three motor types and two circuit board types disassembled at station 2 and station 3, respectively. The demand arrives for a specific type of a component at a station. For instance, if a customer orders a type 1 circuit board, type 1 circuit board inventory at station 3 is checked. If there is no type 1 circuit board, the order is backordered. A type 2 circuit board cannot be used as a substitute for a type 1 circuit board. A conventional WM visits all stations. Following the disassembly at each station, components are tested. The testing times are normally distributed with the means and standard deviations presented in Table 2. A sensor embedded WM visits only the stations which are responsible for the disassembly of functional components and predecessor components of these components. Moreover, following the disassembly of a component, there is no testing due to sensor information on the condition of the component. Another advantage provided by sensors is the identification of component types prior to disassembly. In case of sensor embedded WMs, the component types in a WM can be determined in the beginning of the line. Based on this information, disassembly of WMs involving a specific component type which is demanded more can be prioritized at a station.

Excess product, subassembly and components are disposed of by a small truck with a load volume of 425 cubic feet. Whenever the total volume of the excess product, subassembly and component inventories become equal to the truck volume, the truck loaded with excess inventory is sent to a recycling facility. Any product, subassembly or component inventory which is greater than *maximum inventory level* is assumed to be excess. Component volumes are given in Table 2. The volume of an EOL WM is taken as 20 cubic feet. Multi Kanban System (MKS) developed by Udomsawat and Gupta [10] is used to control the disassembly line.

Table 1. Precedence relationships among the disassembled components.

Part Name	Code	Precedence Relationship	Station
Metal Cover	A	-	1
Circuit Board	B	A	2
Motor	C	A,B	3

Table 2. Specifications for the disassembled components.

Part Name	Code	Testing Time		Volume (cft)	Demanded ?	Number of Different Types	Disposal Classification
		Mean	Std. Dev.				
Metal Cover	A	-	-	1.344	No	-	Steel scrap
Circuit Board	B	5	1	0.015	Yes	3	Waste
Motor	C	10	2	0.137	Yes	2	Waste

3. DESIGN OF EXPERIMENTS STUDY

In order to present a comprehensive and quantitative evaluation of the SEPs on the performance of the WM disassembly line, we compare the case with SEPs against the case without SEPs under different experimental conditions. The factors and factor levels are given in Table 3. A full factorial design with 35 factors requires an extensive number of experiments. Therefore, experiments were designed using OAs which allow for the determination of main effects by making a minimum number of experiments. Specifically, L81 OA was chosen since it requires 81 experiments while accommodating 40 factors with three levels. Discrete Event Simulation (DES) models of both cases were developed using Arena 11 to determine profit value together with various cost and revenue parameters for each experiment. Each DES experiment was carried out for 60480 minutes, the equivalent of six months with one eight hour shift per day.

The following equation presents the formula used in the DES models for the calculation of profit value.

$$Profit = \overbrace{(SR+CR+SCR)}^{\text{Total Revenue}} - \overbrace{(HC+BC+DC+DPC+TC+TPC)}^{\text{Total Cost}} \qquad (1)$$

Where SR is the total revenue generated by the component sales during the simulated time period (STP), CR is the total revenue generated by the collection of EOL WMs during the STP, SCR is the total revenue generated by selling scrap components, HC is the total holding cost of components, EOL WMs and subassemblies during the STP, BC is the total backorder cost of components during the STP, DC is the total disassembly cost during the STP, DPC is the total disposal cost of components, EOL WMs and subassemblies during the STP, TC is the total testing cost during the STP, TPC is the total transportation cost during the STP.

Metal cover and other steel components (e.g., drum, front and side metal frames) are sold as steel scrap. Motor, circuit board and all the other small components are regarded as waste components. In order to determine total weight of small components such as screws, cables, total weight of the main components is multiplied with a *small component weight factor*. In order to calculate the disposal cost of a waste component, the weight is multiplied with the *disposal cost per pound*. Disposal cost for subassemblies and products are calculated by multiplying the total weight of waste components with the *disposal cost per pound*. Disposal cost for subassemblies and products are increased by a factor called *disposal cost increase factor for EOL WMs*. Scrap revenue for metal cover is calculated by multiplying the weight with the *scrap revenue per pound*. In the calculation of scrap revenue for subassemblies and products, total weight of scrap-able components is multiplied with the *scrap revenue per pound.* Scrap revenue for subassemblies and products are decreased by a factor called *scrap revenue decrease factor for EOL WMs*. The time required to retrieve information from the sensors is assumed to be testing time. Duration of this retrieval process is taken as 10 seconds per WM. The operating cost associated with each trip of the truck is assumed to be $50. For each EOL WM, the facility demands a $15 collection fee.

4. RESULTS

Employing the results of DOE studies, various paired-t tests were performed on different performance measures. Table 4 presents the 95% confidence interval, t-value and p-value for each test. According to this table, SEPs achieve statistically significant savings in holding, backorder, disassembly, disposal, testing and transportation costs while they provide statistically significant improvements in total revenue and profit.

Table 3. Factor levels in DOE study.

No.	Factors	Levels		
		1	**2**	**3**
1	Disposal cost increase factor for EOL WMs	0.05	0.10	0.15
2	Scrap revenue decrease factor for EOL WMs	0.05	0.10	0.15
3	Mean demand rate for Circuit Board Type 1 *(components per hour)*	2	4	6
4	Mean demand rate for Circuit Board Type 2 *(components per hour)*	2	4	6
5	Mean demand rate for Circuit Board Type 3 *(components per hour)*	2	4	6
6	Mean demand rate for Motor Type 1*(components per hour)*	2	4	6
7	Mean demand rate for Motor Type 2*(components per hour)*	2	4	6
8	Mean arrival rate of EOL WMs *(products per hour)*	10	20	30
9	Mean disassembly time for station 1 *(minutes)*	0.25	0.50	0.75
10	Mean disassembly time for station 2 *(minutes)*	0.75	1	1.25
11	Mean disassembly time for station 3 *(minutes)*	0.75	1	1.25
12	Backorder cost rate	0.40	0.60	0.80
13	Disassembly cost per minute *($)*	1	2	3
14	Testing cost per minute *($)*	0.40	0.50	0.60
15	Holding cost rate	0.10	0.20	0.30
16	Weight for Metal Cover *(pounds)*	8	12	16
17	Weight for Circuit Board *(pounds)*	0.5	1	1.5
18	Weight for Motor *(pounds)*	5	10	15
19	Weight of other components *(pounds)*	80	100	120
20	Price for Circuit Board Type 1*($)*	20	40	60
21	Price for Circuit Board Type 2*($)*	30	50	70
22	Price for Circuit Board Type 3*($)*	40	60	80
23	Price for Motor Type 1*($)*	50	75	100
24	Price for Motor Type 2*($)*	100	125	150
25	Disposal cost per pound *($)*	0.30	0.40	0.50
26	Scrap revenue per pound *($)*	0.15	0.20	0.25
27	Maximum inventory level	5	10	15
28	Small component weight factor	0.05	0.10	0.15
29	Probability of a non-functional Circuit Board	0.10	0.20	0.30
30	Probability of a non-functional Motor	0.10	0.20	0.30
31	Probability of a missing Circuit Board	0.05	0.10	0.15
32	Probability of a missing Motor	0.05	0.10	0.15
33	Probability of Type 1 Circuit Board	0.20	0.30	0.40
34	Probability of Type 2 Circuit Board	0.20	0.30	0.40
35	Probability of Type 1 Motor	0.40	0.50	0.60

5. CONCLUSIONS

A viable solution approach to deal with the disassembly yield uncertainty is SEPs which involve sensors implanted during the production process. In this study, we analyzed the impact of SEPs on the various performance measures of a WM disassembly line. First, separate DOE studies based on OAs were performed for the cases with and without SEPs. Then paired-t tests were carried out in order to compare two cases for various performance measures. According to the test results, SEPs increase total revenue and profit while providing significant reduction in total cost.

Table 4. Comparison of SEPs against conventional products based on paired-t test results.

Performance Measure	95% Confidence Interval on Mean Difference (Sensor –No Sensor)	t-value	p-value
Holding Cost	(-48.3281, -30.8654)	-9.02	0.000
Backorder Cost	(-40766.6, -27392.4)	-10.14	0.000
Disassembly Cost	(-18926.2, -13308.2)	-11.42	0.000
Disposal Cost	(-44018.5, -36241.0)	-20.54	0.000
Test Cost	(-105743.4, -93996.0)	-33.84	0.000
Transportation Cost	(-3714.70, -3227.89)	-28.38	0.000
Total Cost	(-206132, -181282)	-31.02	0.000
Total Revenue	(187627, 263997)	11.77	0.000
Profit	(372333, 466537)	17.72	0.000

REFERENCES

[1] Ilgin, M. A. and Gupta, S. M., "Analysis of a Kanban Controlled Disassembly Line with Sensor Embedded Products," in *Proceedings of the 2009 Northeast Decision Sciences Institute Conference*, Uncasville, Connecticut, 2009, 555-560.

[2] Ilgin, M. A. and Gupta, S. M., "Environmentally conscious manufacturing and product recovery (ECMPRO): A review of the state of the art," *Journal of Environmental Management*, 2010, 91, 563-591.

[3] Karlsson, B., "A distributed data processing system for industrial recycling," in *Proceedings of 1997 IEEE Instrumentation and Measurement Technology Conference*, Ottowa, Canada, 1997, 197-200

[4] Karlsson, B., "Fuzzy handling of uncertainty in industrial recycling," in *Proceedings of 1998 IEEE Instrumentation and Measurement Technology Conference*, St. Paul, MN, 1998, 832-836.

[5] Klausner, M., Grimm, W. M., and Hendrickson, C., "Reuse of electric motors in consumer products," *Journal of Industrial Ecology*, 1998, 2, 89-102.

[6] Klausner, M., Grimm, W. M., and Horvath, A., "Integrating product takeback and technical service," in *Proceedings of the 1999 IEEE International Symposium on Electronics and the Environment*, Danvers, MA, 1999, 48-53.

[7] Petriu, E. M., Georganas, N. D., Petriu, D. C., Makrakis, D., and Groza, V. Z., "Sensor-based information appliances," *IEEE Instrumentation & Measurement Magazine*, 2000, 3, 31-35.

[8] Scheidt, L. and Shuqiang, Z., "An approach to achieve reusability of electronic modules," in *Proceedings of the IEEE International Symposium on Electronics and the Environment*, San Francisco, CA, 1994, 331-336.

[9] Simon, M., Bee, G., Moore, P., Pu, J.-S., and Xie, C., "Modelling of the life cycle of products with data acquisition features," *Computers in Industry*, 2001, 45, 111-122.

[10] Udomsawat, G. and Gupta, S. M., "Multikanban system for disassembly line," in *Environment conscious manufacturing*, S. M. Gupta and A. J. D. Lambert, Eds. Boca Raton, FL: CRC Press, 2008, 311-330.

[11] Vadde, S., Kamarthi, S., Gupta, S. M., and Zeid, I., "Product life cycle monitoring via embedded sensors," in *Environment conscious manufacturing*, S. M. Gupta and A. J. D. Lambert, Eds. Boca Raton, FL: CRC Press, 2008, 91-103.

A MODELING OF DISASSEMBLY SYSTEMS WITH REVERSE BLOCKING CONSIDERING TRAFFIC AND BUFFER CAPACITIES IN SORTING PROCESS

Tetsuo Yamada*, Tokyo City University, Yokohama, Japan, +81-45-910-2916, tyamada@tcu.ac.jp
Hisashi Yamamoto, Tokyo Metropolitan University, Japan, +81-42-585-8674 yamamoto@cc.tmit.ac.jp
Jin Mitsuhashi, Tokyo City University, Yokohama, Japan
Issei Wakiyama, Tokyo City University, Yokohama, Japan
(*Corresponding author)

ABSTRACT

For environmentally conscious manufacturing [1][2], the sorting process with reverse blocking [3][4] is the first process of inverse manufacturing systems [5] consisting assembly and disassembly [6], and impacts the total productivity. This study models the sorting station in the disassembly system with the reverse blocking considering the traffic and buffer capacities. First, the conditions and capacities are modeled on a sorting queuing system. Next, a transition diagram of the system is drawn, and their stationary state equations are generally formulated. Finally, the system performance is numerically evaluated and compared to [4] in view of blocking probability and throughput.

INTRODUCTION

For environmentally conscious manufacturing [1][2], manufacturers should design efficient disassembly systems [6] for inverse manufacturing [5] consisting assembly and disassembly. However, the sorting in the first process of the disassembly system has the reverse blocking [3][4] and impacts the total productivity of the inverse manufacturing. The fundamental performances of the system were already demonstrated by a queueing analysis [4]. However, arrival/service (traffic) and buffer capacities in at the sorting station are simplified and ignored to overcome the complexities for the equations and states space explosion.

This study models the sorting station in the disassembly system with the reverse blocking considering the traffic and buffer capacities. First, the conditions and capacities are modeled on a sorting queuing system. Next, a transition diagram of the system is drawn, and the stationary state equations of the system are generally formulated. Finally, the system performance is numerically evaluated and compared to [4].

DISASSEMBLY QUEUEING MODEL WITH REVERSE BLOCKING CONSIDERING TRAFFIC AND BUFFER CAPACITIES

This study develops the sorting process with reverse blocking in a disassembly system [4] to consider sorting traffic (system arrival / sorting service) and buffer capacities as shown in Figure 1. The sorting process consists of one sorting station (station 0) and the K succeeding disassembly stations. It is assumed that the collected products from users first arrive at the sorting station according to Poisson Arrival with the mean arrival rate λ. The in-process inventories at the sorting station have infinite or finite buffer capacities. There are K types of arrival products at the system, and the mixed ratio for product i is assumed and set as q_i ($i=1, 2, \ldots, K$, $\sum q_i = 1$). If an arriving product is identified as product i after being serviced at the sorting station, this product is sorted into product i and sent to the succeeding station i. Then, the routing probability from the sorting station to station i becomes q_i. Also, it is assumed that the sorting and each succeeding disassembly station i independently follow the exponential service with the mean service rate μ_i (i=0, 1, …, K). The finite buffer capacity at each succeeding station is set as B_i at station i (i=1,…, K).

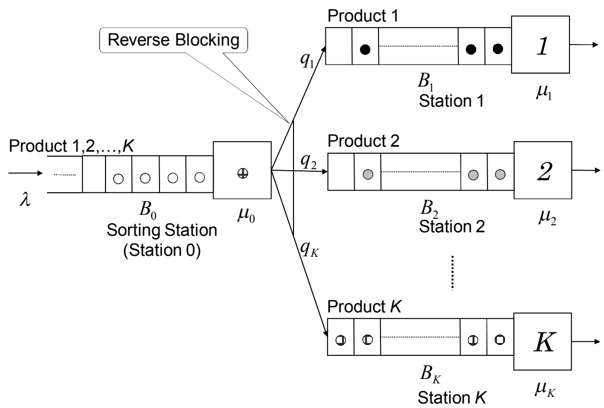


Figure 1 A disassembly queueing model with reverse blocking with sorting traffic/buffer capacities

Unlike [4], this study focuses on not only the sorting but also each succeeding disassembly station for the number of in-process inventories to calculate a system throughput and reverse blocking probability. Each state is represented as state $(l_0, l_1, ..., l_i, ..., l_K)$, where the number of in-process inventories is l_i at the sorting and each succeeding disassembly station i (l_0=0,1, …,B_0, B_0+1, l_i=0,1, …,B_i,B_i+1, B_i+2, i = 1,2, …, K). The reverse blocking phenomenon [3][4] occurs in the queuing model, and is described as follows:

When the buffer capacity at the succeeding station i is full (i.e., the number of in-process inventories, $l_i = B_i + 1$), a new product at the sorting station has just been processed and then identified as a type i product. However, the type i product cannot enter the succeeding disassembly station i because the buffer capacity at station i is already full. Therefore, the product stays at the preceding sorting station, and thus it blocks and stops the sorting station until the service at station i is completed and the buffer capacity becomes available. We represent this reverse blocking state as $l_i = B_i + 2$ (i = 1,2, …, K). It means that one shared buffer space among the succeeding stations is assumed and set at the preceding sorting station during the blocking. Let $P(l_0, l_1, ..., l_i, ..., l_K)$ be the stationary probability for a state $(l_0, l_1, ..., l_i, ..., l_K)$. A summary of the notation used in this paper is given below:

ρ_i	:	utilization rate at station i (i = 0,1, …, K)
l_i	:	number of in-process inventories at station i (i = 0,1, …, K)
BL_i	:	reverse blocking probability at station i (i = 1,2, …, K) ($BL_i = \sum_{(l_0, l_1, ..., l_K) \in C_i} P(l_0, l_1, ..., l_j, ..., l_K)$)
BL	:	total reverse blocking probability at system ($BL = \sum_{i=1}^{K} BL_i$)
TH_i	:	throughput at station i (i =1, …, K) ($TH_i = \mu_i \left\{ 1 - \sum_{(l_0, l_1, ..., l_K) \in A_i} P(l_0, l_1, ..., l_j, ..., l_K) \right\}$)

TH : throughput at system ($TH = \sum_{i=1}^{K} TH_i$)

The assumptions of the model used in this paper are given below:

1) The system is stationary.
2) The travel time of each product is zero.
3) There is no failure at each station.
4) The dispatching rule at all stations is First Come First Served (FCFS).
5) $\rho_0 = \lambda / \mu_0 < 1$, $\rho_i = \mu_0 q_i / \mu_i < 1$ (i =*1*,*2*, …, *K*).

STATIONARY STATE EQUATIONS FOR TRAFFIC AND BUFFER CAPACITIES IN SORTING STATION

In the case of *K* stations with the buffer capacity B_i at station *i*, the stationary state equations for the sorting process are generally formulated such that the sum of output rates from a state in the left side equal the sum of input rates to the state in the right side at each equation, and written as follows:

For $l_0 = 0,1,\ldots,B_0,B_0+1$ and $l_i = 0,1,\ldots,B_i,B_i+1,B_i+2$ $(i = 1,\ldots,K)$,

$$\left(\lambda\beta + \sum_{i=1}^{K} \mu_i \alpha_i + I(x) \sum_{i=1}^{K} \mu_0 q_i \right) P(l_0, l_1, \ldots, l_i, \ldots, l_K)$$

$$= \left\{ \sum_{i=1}^{K} \sum_{x_0 \in \{l_0-1, l_0, l_0+1\}} \sum_{x_i \in \{l_i-1, l_i, l_i+1\}} P(x_0, l_1, \ldots, x_i, \ldots, l_K) - (K-1) \sum_{x_0 \in l_0 - 1} \sum_{x_i \in l_i} P(x_0, l_1, \ldots, x_i, \ldots, l_K) \right\} f_i(x_0, x_i), \quad (1)$$

$$P(l_0, l_1, \ldots, l_i, \ldots, l_j, \ldots, l_K) = 0, \quad \text{if there exist } l_i \text{ and } l_j \text{ such that } l_i = B_i + 2 \text{ and } l_j = B_j + 2, \quad (2)$$

$$\sum_{l_0 \in \{0,1,\cdots,B_0,B_0+1\}} \sum_{\substack{l_i \in \{0,1,\cdots,B_i,B_i+1,B_i+2\} \\ (i=1,\ldots,K)}} P(l_0, l_1, \cdots, l_i, \cdots, l_K) = 1, \quad (3)$$

where

$$\alpha_i = \begin{cases} 0 & \text{if } l_i = 0 \\ 1 & \text{if } l_i \geq 1 \end{cases},$$

$$\beta = \begin{cases} 0 & \text{if } \; l_0 = B_0 + 1 \; \text{ or there exists } l_i \text{ such that } l_i = B_i + 2, l_0 = B_0 \\ 1 & \text{otherwise} \end{cases},$$

$$I(x) = \begin{cases} 0 & x = 0 \\ 1 & x \neq 0 \end{cases}, \qquad x = \left[l_0 \times \prod_{i=1}^{K} \{ l_i - (B_i + 2) \} \right]$$

and $$f_i(x_0, x_i) = \begin{cases} \lambda & \text{if } \; x_0 = l_0 - 1 \; (l_0 \geq 1) \; \text{ and } \; x_i = l_i \\ \mu_i & \text{if } \; x_0 = l_0 \; \text{ and } \; x_i = l_i + 1 \; (i \geq 1) \\ \mu_0 q_i & \text{if } \; x_0 = l_0 + 1 \; \text{ and } \; x_i = l_i - 1 \\ 0 & \text{otherwise} \end{cases}.$$

s.t. $$\sum_{i=1}^{K} q_i = 1. \quad (4)$$

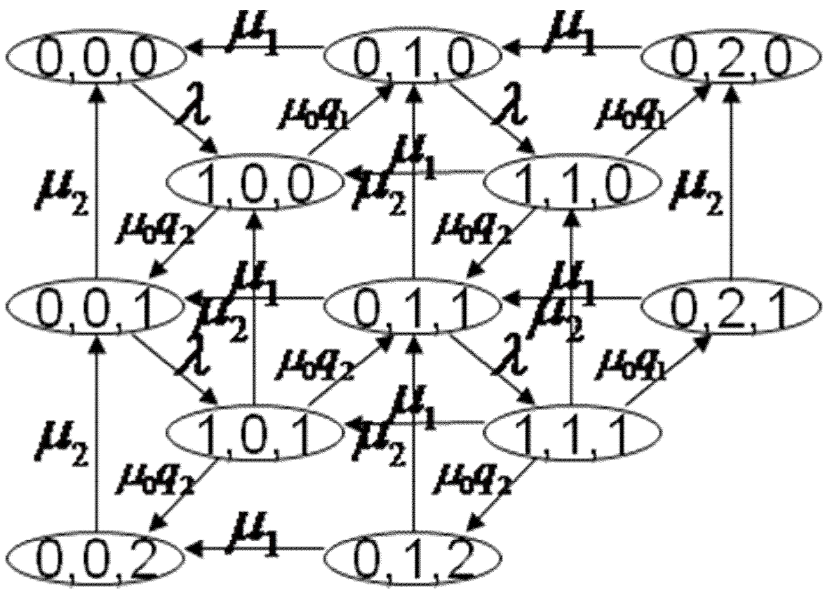


Figure 2 A transition diagram of the states and transitions rates:
Case of 2 stations with finite sorting buffer B_0 ($K=2$, $B_0=B_1=B_2=0$)

Figure 2 shows a transition diagram of the states and transitions rates in the case of 2 stations with the finite sorting buffer B_0 at station i ($i=1,2$). In both right and bottom edges of the Figure 2, the states $(0, 2, l_2)$ for $l_2=1,2$ and $(0, l_1, 2)$ for $l_1=1,2$ become reverse blocking. Its stationary state equations in the cases of infinite and finite sorting buffer capacities are respectively written as follows:

Case of infinite sorting buffer capacity

$$\lambda P(0,0,0) = \mu_1 P(0,1,0) + \mu_2 P(0,0,1)$$
$$(\lambda + \mu_2)P(0,0,1) = \mu_1 P(0,1,1) + \mu_0 q_2 P(1,0,0) + \mu_2 P(0,0,2)$$
$$(\lambda + \mu_1)P(0,1,0) = \mu_2 P(0,1,1) + \mu_1 P(0,2,0) + \mu_0 q_1 P(1,0,0)$$
$$(\lambda + \mu_1 + \mu_2)P(0,1,1) = \mu_1 P(0,2,1) + \mu_2 P(0,1,2)$$
$$+ \mu_0 q_2 P(1,1,0) + \mu_0 q_1 P(1,0,1)$$
$$(\lambda + \mu_2)P(0,0,2) = \mu_1 P(0,1,2) + \mu_0 q_2 P(1,0,1)$$
$$(\lambda + \mu_1)P(0,2,0) = \mu_2 P(0,2,1) + \mu_0 q_1 P(1,1,0)$$
$$(\lambda + \mu_1 + \mu_2)P(0,1,2) = \mu_0 q_2 P(1,1,1)$$
$$(\lambda + \mu_1 + \mu_2)P(0,2,1) = \mu_0 q_1 P(1,1,1)$$

Case of finite sorting buffer capacity

$$\lambda P(0,0,0) = \mu_1 P(0,1,0) + \mu_2 P(0,0,1)$$
$$(\mu_1 + \lambda)P(0,1,0) = \mu_1 P(0,2,0) + \mu_2 P(0,1,1) + \mu_0 q_1 P(1,0,0)$$
$$(\mu_2 + \lambda)P(0,0,1) = \mu_1 P(0,1,1) + \mu_2 P(0,0,2) + \mu_0 q_2 P(1,0,0)$$
$$(\mu_1 + \mu_2 + \lambda)P(0,1,1)$$
$$= \mu_1 P(0,2,1) + \mu_2 P(0,1,2) + \mu_0 q_1 P(1,0,1) + \mu_0 q_2 P(1,1,0)$$
$$\mu_1 P(0,2,0) = \mu_2 P(0,2,1) + \mu_0 q_1 P(1,1,0)$$
$$\mu_2 P(0,0,2) = \mu_1 P(0,1,2) + \mu_0 q_2 P(1,0,1)$$
$$(\mu_1 + \mu_2)P(0,2,1) = \mu_0 q_1 P(1,1,1)$$
$$(\mu_1 + \mu_2)P(0,1,2) = \mu_0 q_2 P(1,1,1)$$
$$(\mu_0 q_1 + \mu_0 q_2)P(1,0,0) = \mu_1 P(1,1,0) + \mu_2 P(1,0,1) + \lambda P(0,0,0)$$
$$(\mu_1 + \mu_0 q_1 + \mu_0 q_2)P(1,1,0) = \mu_2 P(1,1,1) + \lambda P(0,1,0)$$
$$(\mu_2 + \mu_0 q_1 + \mu_0 q_2)P(1,0,1) = \mu_1 P(1,1,1) + \lambda P(0,0,1)$$
$$(\mu_1 + \mu_2 + \mu_0 q_1 + \mu_0 q_2)P(1,1,1) = \lambda P(0,1,1)$$

EFFECT OF SORTING TRAFFIC AND BUFFER CAPACITIES

In this section, the fundamental characteristics for the throughput and the reverse blocking are discussed when the sorting traffic (system arrival / service rate) and buffer capacities is considered. The numerical experiments are conducted using numerical analysis software, Maple 11, to solve the difficult stationary state equations. The performance by the system arrival rate λ, the service rate μ_i and the buffer capacity B_i are numerically considered in the case of 2 stations as follows:

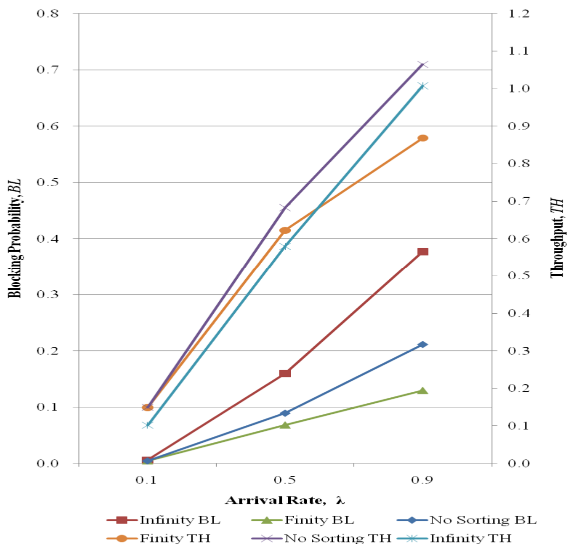


Figure 3 Behaviors of blocking probability *BL* and throughput *TH* for system arrival rate λ (K=2, q_1=q_2=0.5, $\mu_0=1$, $\mu_1=\mu_2=1$, B_1=B_2=0,)

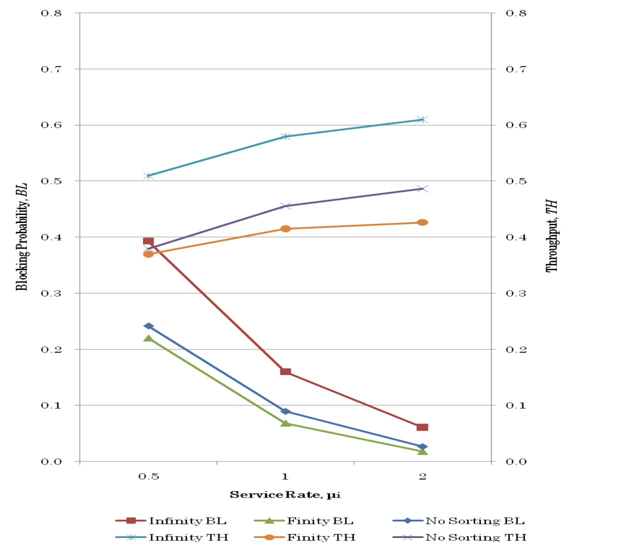


Figure 4 Behaviors of blocking probability *BL* and throughput *TH* for disassembly service rate μ_i (K=2, $\lambda=0.5$, q_1=q_2=0.5, $\mu_1=\mu_2$, B_1=B_2=0)

As the system arrival rate λ increases (Figure 3), the total reverse blocking probability *BL* increases, however, the system throughput *TH* also increases in the both cases of infinite and finite sorting buffers as well as the case of no sorting capacity. As the disassembly service capacity μ_i (i=1,2) increases (Figure 4), the blocking probability *BL* decreases, and then the throughput *TH* increases for the all 3 cases. Therefore, it is seen that the blocking and throughput behaviors in the case of no sorting capacity [4] can be used for the system arrival rate and the disassembly service rate.

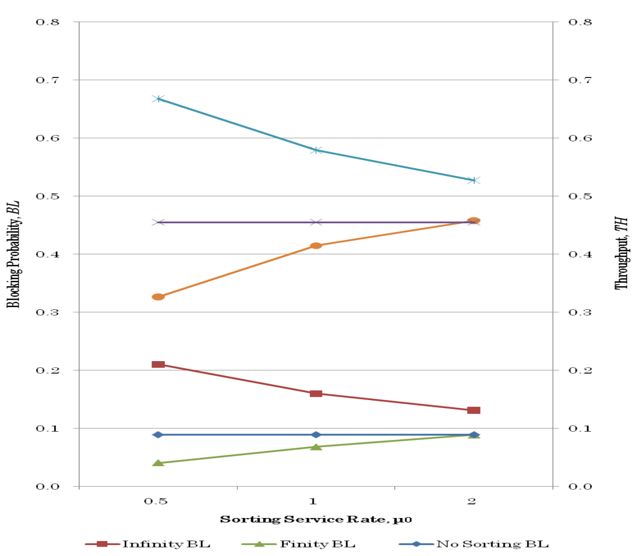


Figure 5 Behaviors of blocking probability *BL* and throughput *TH* for sorting service rate μ_0 (K=2, $\lambda=0.5$, q_1=q_2=0.5, $\mu_1=\mu_2=1$, B_1=B_2=0)

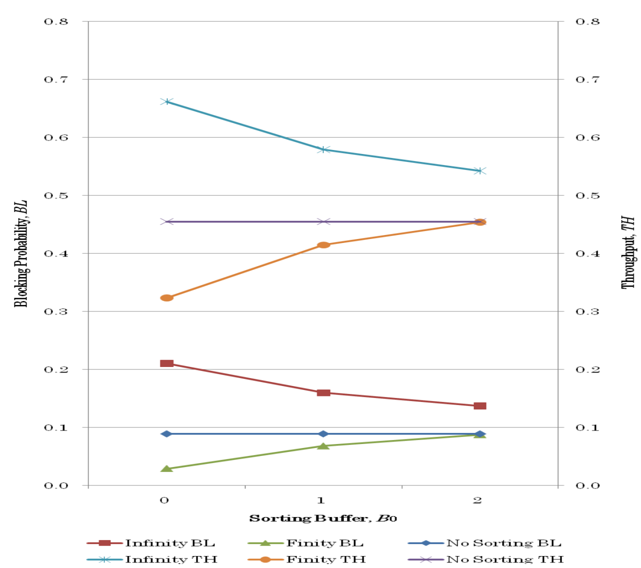


Figure 6 Behaviors of blocking probability *BL* and throughput *TH* for sorting buffer capacity B_0 (K=2, $\lambda=0.5$, q_1=q_2=0.5, B_1=B_2=0, $\mu_1=\mu_2=1$)

The effectiveness of sorting service and buffer capacities are here considered. As the sorting service rate μ_0 increases (Figure 5), the blocking probability *BL* decreases in the case of the infinite sorting buffer case, but decreases in the case of the finite one. Finally, the values in the both cases reach to *BL*=0.09 in the case of no sorting capacity. Similar to the blocking probability *BL*, the throughput *TH* decreases in the case of the infinite sorting buffer case, but increases in the case of the finite one. Finally, the values in the both cases reach to *TH*=0.46 in the case of no sorting capacity. Also, the behaviors for the sorting buffer capacity B_0 in Figure 6 is similar to one for the sorting service rate in Figure 5. Therefore, the sorting service/buffer designs in this study should be considered when those capacities are lower.

SUMMARY AND FUTURE STUDIES

This study modeled the disassembly system with the reverse blocking considering the traffic (arrival/service) and buffer capacities in the sorting station. First, the conditions and capacities were modeled on a sorting queuing system. Next, a transition diagram of the system were drawn, and the stationary state equations of the system were formulated in the case of 2 stations. Finally, the system performance is numerically evaluated and compared to [4].

For the system arrival and the disassembly service rates, it is seen that the blocking and throughput behaviors in the cases of infinite/finite sorting capacity are similar to ones in the case of no sorting capacity [4]. Also, for the sorting service rate and buffer capacities, the reverse blocking probability and the system throughput inversely behave in the cases of infinite and finite buffer capacities, and the values in the both cases reach to ones in the case of no sorting capacity [4]. Therefore, the sorting service/buffer designs in this study should be considered when those capacities are lower.

Future studies should evaluate balances among arrival, sorting and disassembly service capacities in this model, design a lager disassembly system with reverse blocking than one of this study, and model the entire disassembly and/or inverse manufacturing systems, etc.

REFERENCES

[1] Ilgin, M.A., and Gupta, S. M., "Environmentally Conscious Manufacturing and Product Recovery: A Review of the State of the Art", *Journal of Environmental Management*, 2010, 91(3), 563-591.

[2] Gupta, S. M., Lambert, A. J. D., *Environment Conscious Manufacturing*. Boca Raton, FL: CRC Press, 2008.

[3] Yamada, T., Matsui, M.. "Disassembly Production Systems and Its Design Issues", *Reprints of Japan Industrial Management Association, Spring Meeting*, 2003, Chofu, Japan, May, 144–145 (in Japanese).

[4] Yamada, T., Mizuhara, N., Yamamoto, H., Matsui, M., "A Performance Evaluation of Disassembly Systems with Reverse Blocking", *Computers & Industrial Engineering*, 2009, 56(3), 1113-1125.

[5] Inverse Manufacturing Forum, *Handbook of Inverse Manufacturing*. Tokyo, Japan: Maruzen, 2004 (in Japanese).

[6] Lambert, A. J. D., Gupta, S. M., *Disassembly Modeling for Assembly, Maintenance, Reuse, and Recycling*. Boca Raton, FL: CRC Press, 2005.

AN EOQ MODEL WITH TWO TYPES OF SHORTAGE AND RANDOM DEFECTIVE UNITS

Javad Paknejad, Hofstra University, (516)463-4244, M.J.Paknejad@hofstra.edu
Farrokh Nasri, Hofstra University, (516)463-4244, Farrokh.Nasri@hofstra.edu
John F. Affisco, Hofstra University, (516)463-4244, John.F.Affisco@hofstra.edu

ABSTRACT

This paper extends the results developed by Nasri, Paknejad, and Affisco [7] to the case of basic Economic Order Quantity model with random defective units and two types of shortage. In addition to the general relationships for the optimal values of policy variables, the paper also provides closed form expressions when proportion of defective items in a lot follows a one-sided truncated exponential distribution.

Keywords: Inventory Theory, Quality

INTRODUCTION

Quality movements have brought the importance of small lot sizes of high quality to the forefront. In his seminal work Schonberger [13] indicates that small lots of high quality lead to decreased costs due to minimization of discard, waste, and rework; improved worker motivation; decreased buffer stocks; and reduced cycle inventories. These in turn, result in increased productivity and improved market response. This set of principles is reinforced in Hall [5] where a theme of continuous improvement aimed at achieving reduced lot sizes through improved quality is fostered.

Realizing the importance of quality in operations, researchers have developed significant number of inventory models that study the relationships among lot size and quality. Initially, Rosenblatt and Lee [12] investigated the effect of process quality on lot size in the classical Economic Manufacturing Quantity (EMQ) model. Porteus [11] introduced a modified EOQ model that indicates a significant relationship between quality and lot size. In both these cases, demand is assumed to be deterministic. Moinzadeh and Lee [6] investigated the effect of defective items on the order quantity and reorder point of a continuous-review inventory model with Poisson demand and constant lead time. Paknejad, Nasri, and Affisco [10] extend this work to consider stochastic demand and constant lead time in the continuous review (s,Q) model. They [9] also develop a quality-adjusted EOQ model for the case where both backorders and stockouts are allowed.

Cheng [3] develops a model that integrates quality considerations with the Economic Production Quantity (EPQ). The author assumes that unit production cost increases with increases in process capability and quality assurance expenses. Classical optimization results in closed forms for the optimal lot size and optimal expected fraction acceptable. The optimal lot size is intuitively appealing since it indicates an inverse relationship between lot size and process capability. It should be noted that a good survey of the early literature on integrating lot size and quality control policies is given in Goyal et al [4].

In this early work the authors mainly assume that the manufacturer operates a process that is in statistical control. That is, the process generates a known, constant proportion of defectives, θ. Such an assumption induces a situation where the proportion of defective items follows a binomial distribution, and process quality, therefore, may be monitored by a proportion control chart. This assumption is also made in Affisco, Paknejad, and Nasri [1] for the case of the EOQ and Affisco, Paknejad, and Nasri [2] for the case of the joint economic lot size model.

In two recent papers, Nasri, Paknejad, and Affisco [7,8] studied the relationship between order quantity and quality for processes that have not yet achieved the state of statistical control. Specifically, the authors in [7,8] considered the case of the EOQ and EMQ models with backorders where the number of defectives produced by the manufacturing process is random, rather than constant, and adjusted the EOQ and EMQ models, respectively, for the quality factor. In addition to general relationships obtained, Nasri, Paknejad, and Affisco [7,8] also considered special cases of proportion of non-defectives in each lot and presented explicit results for cases considered.

One of the assumptions of the models developed in [7,8] for the purpose of this paper is that the shortage cost is calculated based on the number of shortages per year. However, it may sometimes be more appropriate to assess the unit shortage cost against the average number of shortages irrespective of the duration of the shortage. In the real world, it is not always easy to argue which of the two cases of shortage cost apply. Hence, in this paper, we extend the previous work in [7,8] by combining two alternative assumptions commonly made about the unit shortage cost.

MODEL AND ASSUMPTIONS

Consider the basic Economic Order Quantity model that allows stockouts and backordering with two distinct types of shortage and the following total annual cost function

$$TC(S,Q) = \frac{D}{Q}K + \frac{(Q-S)^2}{2Q}c_h + \frac{S^2}{2Q}c_b + \frac{SD}{Q}c_b' \qquad (1)$$

where

D = Annual Demand in units,

Q = Lot size per order,

S = Number of units backordered,

K = Setup cost per setup,

c_h = Holding cost per unit per year,

c_b = Shortage cost per unit short per year,
c_b' = Shortage cost per unit short, irrespective of the duration of shortage.

The results of classical optimization yields the following well-known expressions for the optimal values for the lot size, Q^*, and units backordered, S^*, are as follows

$$Q^* = \sqrt{\left[\frac{2DK}{c_h} - \frac{\left(Dc_b'\right)^2}{c_h\,(c_h + c_b)}\right]\left(\frac{c_h + c_b}{c_b}\right)} \tag{2}$$

$$S^* = Q^*\left(\frac{c_h}{c_h + c_b}\right) - D\left(\frac{c_b'}{c_h + c_b}\right) \tag{3}$$

Implicit in these derivations is that all units produced by the vendor, in response to the purchaser's order, are of acceptable quality. Now, assume that this is not the case. Specifically, assume that each lot contains a random proportion of defective units. Upon arrival, the purchaser inspects the entire lot. We further assume that the purchaser's inspection process is perfect, and that all rejected items are returned to the vendor at no cost to the purchaser. In addition, we assume that the inspection cost is paid by the vendor. Of course, it is likely that the vendor will recover some of these costs from the purchaser either directly or indirectly. Based on this scenario, we now adjust the EOQ with planned shortage model for two types of shortage and the quality factor.

Let
θ = Proportion of defective items in a lot, a continuous random variable,
$f(\theta)$ = Probability density function of θ,
$y = (1-\theta)Q$ = Number of non-defective items in a lot,
$c(y)$ = Cost per cycle given that there are y non-defective items in the lot,
$T = y/D$ = Cycle time (time between two successive placement of orders),
$E(y)$ = First moment of y, given a lot of size Q is ordered,
$E(y^2)$ = Second moment of y given a lot of size Q is ordered,
$E(c)$ = Expected cycle cost per year,
$E(T)$ = Expected value of T given a lot of size Q is ordered,
$C(S,Q)$ = Expected total cost per year.

The total cost per cycle is

$$c(y) = K + \frac{[(1-\theta)Q - S]^2}{2D}c_h + \frac{S^2}{2D}c_b + S\,c_b' \tag{4}$$

The average cycle time and cycle cost are

$$E(T) = \frac{E(y)}{D} = \frac{E[(1-\theta)\,Q]}{D} = \frac{Q[1-E(\theta)]}{D} \tag{5}$$

and

$$E(c) = K + \frac{c_h.Q}{2D}\left\{Q\left[1-2E(\theta)+E(\theta^2)\right]-2S\left[1-E(\theta)\right]\right\} + \frac{c_h + c_b}{2D}.S^2 + c_b^{'}.S \tag{6}$$

The expected total annual cost is

$$c(S,Q) = \frac{DK}{Q\left[1-E(\theta)\right]} + \left[\frac{Q\left[1-2E(\theta)+E(\theta^2)\right]}{2\left[1-E(\theta)\right]}\right]c_h + \frac{(c_h + c_b)S^2}{2Q\left[1-E(\theta)\right]} + \left[\frac{c_b^{'}.D}{Q\left[1-E(\theta)\right]} - c_h\right].S \tag{7}$$

The optimal values for the order quantity, Q_{adj}^*, units backordered, S_{adj}^*, are found by using calculus as follows

$$Q_{adj}^{*} = \frac{1}{1-E(\theta)}\sqrt{\frac{2DK - \frac{c_b^{'2}}{c_h + c_b}D^2}{c_h\left[\frac{1-2E(\theta)+E(\theta^2)}{\left[1-E(\theta)\right]^2} - \frac{c_h}{c_h + c_b}\right]}} \tag{8}$$

$$S_{adj}^* = \left[1 - E(\theta)\right]\left(\frac{c_h}{c_h + c_b}\right)Q_{adj}^* - \frac{c_b^{'}}{c_h + c_b}D \tag{9}$$

Note that in equations (8) and (9) when $c_b^{'}=0$, the results of this paper simply reduce to the corresponding results in Nasri, Paknejad, and Affisco [7]. Furthermore, if θ=0, then quality is perfect and equations (8) and (9) become identical to equations (2) and (3) of the basic EOQ with two types of shortage model.

However, when $c_b > 0$ and $c_b^{'} > 0$, it is possible to find a negative value under the radical in (8). In this case, the optimal policy is to permit no shortages (S_{adj}^*=0), which results in a lot size given by

$$Q_{adj}^* = \sqrt{\frac{2DK}{c_h\left[1-2E(\theta)+E(\theta^2)\right]}} \tag{10}$$

Finally, when c_b=0 and $c_b^{'} > 0$, the optimal policy is either to allow no shortages or to not order the item from the supplier until a sufficient backlog is accumulated.

Let us now consider a special case of θ as a continuous random variable. Specifically, let θ be a one sided Truncated Exponential, TEXP(1/λ,1), random variable. In this case the probability density function of θ and its first and second moments are as follows

$$f_\theta(\theta \mid 1/\lambda, 1) = \frac{\lambda e^{-\theta\lambda}}{1-e^{-\lambda}} \qquad for \quad 0 < \theta \le 1 \quad where \quad \lambda > 0 \tag{11}$$

$$E(\theta) = \frac{1}{\lambda}\left[\frac{(1-e^{-\lambda}) - \lambda e^{-\lambda}}{1-e^{-\lambda}}\right] \tag{12}$$

and

$$E(\theta^2) = \frac{2}{\lambda^2}\left[\frac{1-\frac{1}{2}(\lambda^2+2\lambda+2)e^{-\lambda}}{1-e^{-\lambda}}\right] \tag{13}$$

In this case

$$Q^*_{adj,TE} = \frac{\lambda\left(1-e^{-\lambda}\right)}{\lambda-\left(1-e^{-\lambda}\right)}\sqrt{\frac{2DK - \frac{c_b'^2}{c_h+c_b}D^2}{c_h\left[\frac{\left(1-e^{-\lambda}\right)\left\{\lambda^2-2\lambda+2-2e^{-\lambda}\right\}}{\left(\lambda-1+e^{-\lambda}\right)^2} - \frac{c_h}{c_h+c_b}\right]}} \tag{14}$$

and

$$S^*_{adj,TE} = \left(\frac{1}{1-e^{-\lambda}} - \frac{1}{\lambda}\right)\left(\frac{c_h}{c_h+c_b}\right)Q^*_{adj} - \frac{c_b'}{c_h+c_b}D \tag{15}$$

Similar results can easily be found for other forms of probability density function of θ such as Triangular, Beta, etc.

CONCLUSION

Recently, in two related papers, Nasri, Paknejad, and Affisco [7,8] opened a new line of inquiry into EOQ and EMQ models with planned shortage subject to imperfect quality. Assuming that shortage cost is levied based on per unit short per year, closed forms were derived for the cases of infinite and finite production rates, respectively, where the manufacturing process is unstable, with no known process capability, and the proportion of defective items is randomly distributed.

This paper extended the infinite production rate model developed by Nasri, Paknejad, and Affisco [7] by combining two distinct types of shortage cost in the model. In addition to the general relationships, closed forms are also given for the special case where the proportion of defective items follows a one-sided truncated exponential density function.

Acknowledgement: This research was partially supported by a Summer Research Grant from the Frank G. Zarb School of Business.

REFERENCES

[1] Affisco, J.F., Paknejad, M.J., and Nasri, F. "Quality improvement in the economic order quantity model," *Proceedings of the Northeast Decision Sciences Institute*, pp. 256-258, Atlantic City, New Jersey, 2000.

[2] Affisco, J.F., Paknejad, M.J., and Nasri, F. "Quality improvement and setup reduction in the joint economic lot size model, "*European Journal of perational Research*," 2002, 142, 497-508.

[3] Cheng, T.C.E., 1991, "EPQ with process capability and quality assurance considerations," *Journal of Operational Research Society*, 42(8), 713-720.

[4] Goyal, S.K., Gunasekaran, A., Martikainen, T. & Yli-Olli, P., 1993, "Integrating production and quality control policies: A Survey," *European Journal of Operational Research*, 69, 1-13.

[5] Hall, R.W. *Zero Inventories.* Dow Jones-Irwin, Homewood, IL, 1983.

[6] Moinzadeh, K., and H.L., Lee, "A continuous review inventory model with constant resupply time and defective items," *Naval Research Logistics*, 1987, 34, 457-467.

[7] Nasri, F., Paknejad, M.J., & Affisco, J.F., "Lot size determination in an EOQ model with planned shortage and random defective units," *Proceedings of the Northeast Decision Sciences Institute*, pp. 403-408, Mohegan Sun, CT, 2009.

[8] Nasri, F., Paknejad, M.J., & Affisco, J.F., "An EMQ model with planned shortage and random defective units," *Proceedings of the Decision Sciences Institute*, pp. 10541-10546, New Orleans, 2009.

[9] Paknejad, M.J., Nasri, F. & Affisco, J.F., "Flexibility and Quality Improvement in the EOQ with Planned Shortage Model," *Proceedings of the Northeast Decision Sciences Institute*, pp. 257-259, Pittsburgh, Pennsylvania, 2001.

[10] Paknejad, M.J., Nasri, F. & Affisco, J.F., "Defective units in a continuous review (s,Q) system," *International Journal of Production Research*, 1995, 33(10), 2767-2777.

[11] Porteus, E., "Optimal lot sizing, process quality improvement and setup cost reduction," *Operations Research*, 1986, 34(1), 137-144.

[12] Rosenblatt, M.J. and Lee, H.L., "Economic production cycles with imperfect production processes," *IIE Transactions*, 18(3), 1986, 48-55.

[13] Shonberger, R.A., *Japanese manufacturing techniques – Nine hidden lessons in simplicity*. The Free Press, New York, 1982.

SYSTEM DYNAMICS APPROACH IN INVENTORY MANAGEMENT OF EOL PRODUCTS IN A DISASSEMBLY LINE

Badr O. Johar, Northeastern University, (617) 373-7635, johar.b@neu.edu
Surendra M. Gupta, Northeastern University, (617) 373-4846, gupta@neu.edu

ABSTRACT

A major challenge original equipment manufacturers (OEMs) currently face is how to implement an effective reverse supply chain network that is both cost effective and efficient. The rapid increase in the end-of-life (EOL) products returns from end customers back to origin for recovery or proper disposal can be said one of the main reasons behind this interest. Thus, the purpose of this paper is to examine the relationship between all elements of inventory management in the disassembly context using the system dynamic modeling approach to analyze the inventory of disassembled parts. Analysis of different market scenarios is conducted. The system dynamic tool can be used as an experimental tool which can be used to forecast the system's behavior under different conditions.

Keywords: Reverse supply chain, Disassembly, System dynamic, End-of-life, Inventory management

INTRODUCTION

Recent years have witnessed Reverse Supply Chain (RSC) become the center of attention for researchers and Original Equipment Manufacturers (OEM's). “Reverse supply chain is the series of activities required to retrieve a used product from a customer and either dispose of it or reuse it” [1]. The fast depletion of virgin resources and the rapid increase of product returns from end customers to original manufacturers for maintenance, repair or to be disposed of can be said to be one of the main reasons behind this interest. Electronics manufacturers have introduced state-of-the-art technologies in quick succession. As a result of this, the end-of-life (EOL) products returned by customers have grown significantly in volume. But often these end-of-life (EOL) products are found to be in excellent working conditions (functional). Customers return them because of the various marketing programs or favorable incentives offered by service providers or manufacturers that create a “must have” sense in the minds of customers to acquire upgraded products. Of course, there are times when the products are worn out or do not function properly anymore. In the past two decades environmental concern has focused on production processes, and environmental regulation has concentrated on pollution from industry. However, there is growing awareness that this may not be sufficient and it is increasingly recognized that the use and disposal phases, as well as the production phase of the product life cycle, are important. Environmentalists have always demanded that the manufacturing companies should take these EOL products back and manage them in an environmentally conscious manner. EOL products can be remanufactured, reused, recycled, or disposed of [2], [3]. However, planners did not invest or engage themselves in such initiatives because of the uncertainty in quantity and quality of items taken back.

BACKGROUND IN DISASSEMBLY

Environmental regulations and consumer pressure are forcing manufacturers to become more responsible for the safe disposal and recycling of used products. This requires a new approach to product design, one which results in a product designed for all the stages of its life-cycle. Many corporations have understood the economic and environmental benefits of minimizing the use of virgin resources. Also, due to strict environmental legislations, they have started to comprehend the importance of the recovery process and are taking serious steps in restructuring their supply chain processes to meet the new regulations. Examples of such regulations include limitations on waste disposal and recycling requirements. The idea behind the change is to use materials and parts more than once before they are finally discarded. Thus, effective supply chain management is vital in gaining a competitive edge over other corporations. Supply chain management, by definition, “involves functions such as production, purchasing, material management, warehousing, inventory control, distribution, shipping, and transport logistics” [4]. Supply Chain Management is a consequence of the increased necessity for holistic considerations in between and across companies' business activities and resources in and between marketing channels, in order to improve the overall performance towards the ultimate consumer in the marketplace [5].

However, the take-back process is more environment friendly than the traditional supply chain process as it “closes the loop” of the supply chain process and transforms the end-of-life (EOL) products into new serviceable products. This new portion of the supply chain is known as Reverse Supply Chain (RSC). It has been found that the original supplier is in the best position to control the return process. The basic reverse supply chain logistics model operates independently of the forward supply chain that delivered the original product. When a firm controls the full process of forward and backward shipment the result is called a Closed Loop Supply Chain [6], [7].

Closed-Loop Supply Chain Management may be defined as the design, control, and operation of a system to maximize value creation over the entire life cycle of a product with dynamic recovery of value from different types and volumes of returns over time [1]. In comparison to regular supply chain management, the management of reverse supply chain is more challenging because it is more reactive and less visible. Such challenges may include, changing demand rates, multiple demand sources, variety of products and logistical complications. Yet another challenge is inventory control and value management of end-of-life (EOL) products. Due to the disparity between demand for parts and materials and their yields, the planner faces economical as well as physical constraints when trying to take a decision on how many products to take back and when to take those back. Economical constraints may include holding and transaction cost of excess inventory whereas physical constraints may include capacity and space limitations. The objective of this research is to develop and examine an inventory control policy of an On-Hand Inventory (OHI) of returned items and disassembled parts using system dynamic (SD) approach that works as a simulation tool to examine the behavior of the inventory pile up along the disassembly line and the effect of different changes in supply/demand trends on the overall inventory cost.

SYSTEM DYANMIC: METHODOLOGY AND APPROACH

In this paper, system dynamics approach is presented in an attempt to solve the inventory control problem in context of disassembly. The model assumes a stochastic formulation that takes into consideration the uncertainty in demand and line yields. Generally, system dynamics approach is applied to macro-systems such as production-inventory system, national economies, and macroeconomic systems. The philosophy behind it is founded on the concept that the system state changes based on the changes in the rates of inflow and outflow, and characterized by the feedback loop that triggers the corrective action, and thus, the model can be developed using Euler's first order differential equations. This study provides knowledge that helps understanding the challenges and opportunities associated with inventory control in disassembly line context. Disassembly Line faces a serious inventory problem because of the disparity between the demand and the disassembly line yields. Workstation along the line tend to experience different accumulation rates and depletion rates depending on how demanded these parts are. This fluctuation between inventory and demand will create uncertainty in inventory levels.

Methodology

The problem statement clearly indicates that it is extremely difficult to solve this problem without making a number of generalizations. Simulation appears to be the most appropriate solution procedure. However the accuracy of simulation warrants precise and voluminous real life data which is very difficult to obtain.

Hence the author feels that two alternative techniques could be employed:

1) System Dynamics simulation modeling
2) Markov chain analysis

This paper will address the System Dynamic (SD) approach only.

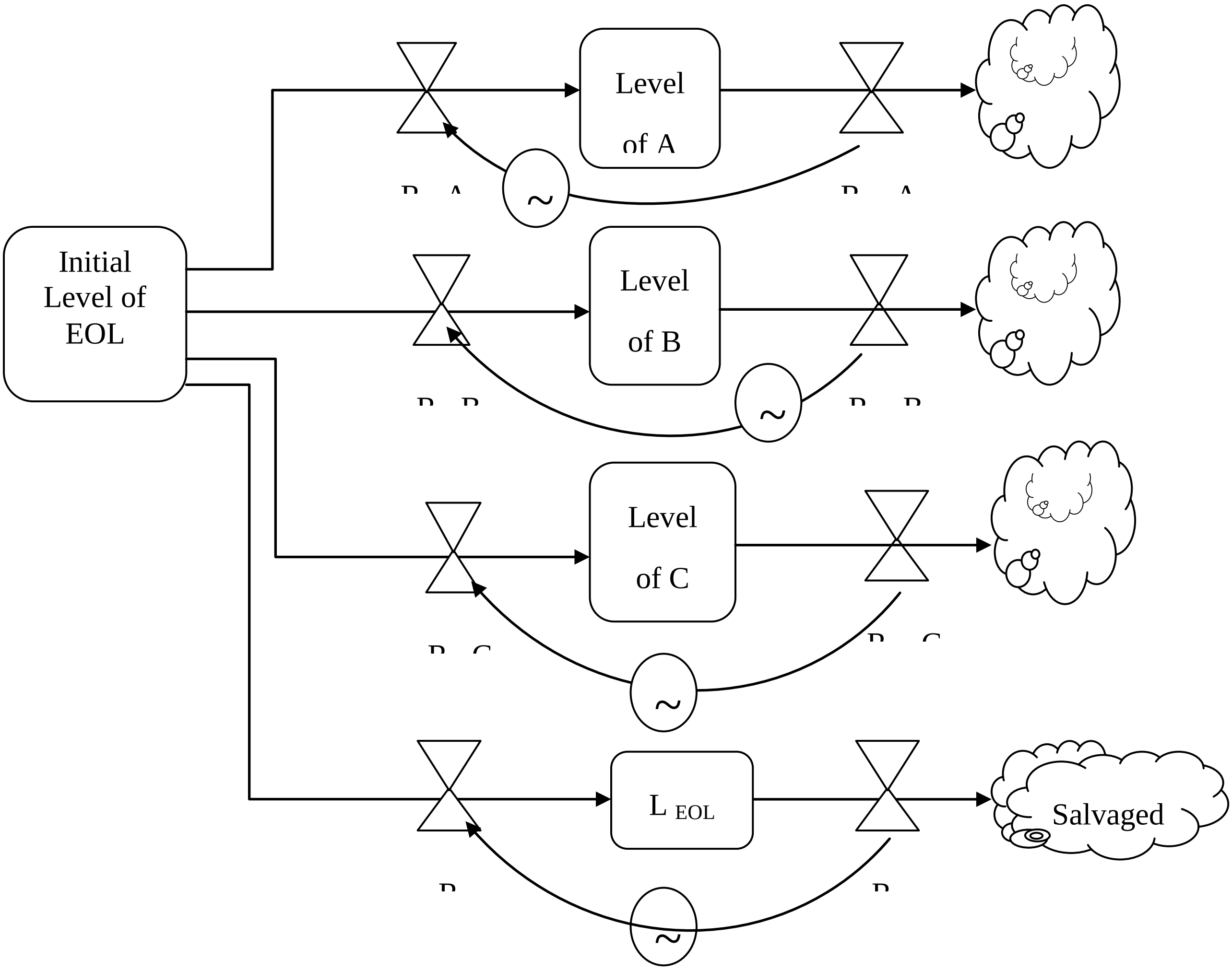
Approach

System Dynamics (SD) approach is generally applied to macro-systems such as production-inventory systems, macroeconomic systems, demographic transformations, national economies and so on. Growth models for national economies have been successfully developed by many researchers. Production-inventory systems were elaborately discussed by Forrester J.W in his ‘Industrial Dynamics’.

The philosophy behind SD simulation is founded on the concept that industrial systems are like input-output systems. The system state changes according to the changes in the rates of inflow and outflow. The steady state condition is achieved when the input rate is equal to the output rate. Thus any system can be viewed as an input-output system with input and output rates. The system is characterized by the feedback loop that triggers the corrective (control) action.

Based on these assumptions the system model can be developed using Euler’s first order difference equations. The model is as shown in figure 1 below:

Figure (1) System Dynamic in Disassembly Line



Notation

L_{EOL} Level of end-of-life (EOL) inventory of core products

L_j Level of disassembled part type *j*, where $j = \{A, B, C\}$

$L_{EOL(t-1)}$ Level of end-of-life (EOL) inventory of core products at time $t-1$

$R_{in}EOL$ Rate of incoming end-of-life (EOL) products, known as input rate

$R_{out}EOL$ Rate of demand of end-of-life (EOL) products, known as output rate

Δt Incremental time or time step

$L_j(t)$ Inventory level of disassembled part *A* at time t

$L_j(t+\Delta t)$ Inventory Level of disassembled part *j*, at time $t+\Delta t$

Model

If the rates of inflow and outflow are precisely known which may depend upon the demand distribution of each product A, B or C, it would be easier to solve the problem. It is also important that the unit costs of each product C_A, C_B, and C_C also known. Similarly an additional complexity that emerges in the modeling is the life period of products A, B or C. Since all these products are supposed to have reached their life period, reusability depends on further deterioration or decay. If an exponential deterioration is assumed the model has to be incorporated with these terms. Here the author intends to solve the problem assuming that deterioration doesn't make any problem in salvaging the product. The author wishes to use SD software ARENA for simulating this system. The problem can be modeled as follow

$$EOL\_INV\_Level = L_{EOL} = L_A + L_B + L_C = L_{EOL(t-1)} + (R_{out}EOL - R_{in}EOL)\Delta t \tag{1}$$

Similarly,

$$L_A\_a\_t + \Delta t = L_A(t+\Delta t) = L_A t + (R_{out}A - R_{in}A)\Delta A \tag{2}$$

$$L_B\_a\_t + \Delta t = L_B(t+\Delta t) = L_B t + (R_{out}B - R_{in}B)\Delta t \tag{3}$$

$$L_C\_a\_t + \Delta t = L_C(t+\Delta t) = L_C t + (R_{out}C - R_{in}C)\Delta t \tag{4}$$

CONCLUSIONS

In this paper, a brief description of the inventory control problem in disassembly line is given. The probabilistic nature of the problem is assumed to provide more accurate results compared to deterministic models, hence systems dynamic (SD) approach is developed for a hypothetical product *ABC*. Problem methodology and model is presented in this paper, and simulation model using commercial software ARENA will be developed to simulate the buildup of inventory and steady state of the system.

REFERENCES

[1] Guide, V. Daniel R. Jr. and van Wassenhove, L. N., "The Evolution of Closed-Loop Supply Chain Research", *Operations Research*, Vol. 57, No. 1, 10-18, 2009.

[2] Gungor, A. and Gupta, S. M., "Issues in Environmentally Conscious Manufacturing and Product Recovery: A Survey", *Computers and Industrial Engineering*, Vol. 36, No. 4, 811-853, 1999.

[3] Ilgin, M. A. and Gupta, S. M., "Environmentally Conscious Manufacturing and Product Recovery (ECMPRO): A Review of the State of the Art", *Journal of Environmental Management*, Vol. 91, No. 3, 563-591, 2010.

[4] Jaber, M. Y., Rosen, M. A., "The Economic Order Quantity Repair and Waste Disposal Model with Entropy Cost", *European Journal of Operational Research*, Vol. 188, No. 1, 109-120, 2008.

[5] Svensson, G., "The Theoretical Foundation of Supply Chain Management: A Functionalist Theory of Marketing", *International Journal of Physical Distribution and Logistics Management*, Vol. 32, No. 9, 734-754, 2002.

[6] Kumar, S., Yamoka, T., "System Dynamics Study of the Japanese Automotive Industry Closed Loop Supply Chain", *Journal of Manufacturing Technology Management*, Vol. 18, No. 2, 115-138, 2007.

[7] Pochampally, K. K., Nukala, S. and Gupta, S. M., *Strategic Planning Models for Reverse and Closed-Loop Supply Chains*, CRC Press, Boca Raton, Florida, ISBN: 9781420054781, 2009.

ASSIGNING PASTORS TO PARISHES: AN OPTIMIZATION MODEL

Keith A. Willoughby, University of Saskatchewan, Saskatoon, SK, Canada S7H 3A8,
(306) 966-2128, willoughby@edwards.usask.ca

ABSTRACT

Pastors in a particular religious diocese in Canada are required to make monthly visits to parishes. Previously, each year's visit schedule was manually constructed. We report on the development of an analytical model to optimize the assignments. Through the use of our approach, the diocese has reduced its annual travel reimbursement costs. It is planning to make permanent use of this optimization model in identifying monthly visit schedules.

KEYWORDS: Mathematical programming/ Optimization

1.0 INTRODUCTION

Decision-makers use assignment models to optimize resource allocation. A variety of successful applications have been reported in the literature, including buyer-seller meetings to trade show appointment periods [3], flight training time blocks to aeronautical students [2], female collegiate students to sorority organizations [4] and medical residents to hospitals [5].

In this paper, we develop an integer optimization model to make cost-effective assignments in a non-profit context. Specifically, we assign monthly visits of religious pastors to parishes. The religious organization explored in this paper includes a large diocese in Saskatchewan, Canada with ten respective parishes (also called units). Figure 1 shows a map of the diocese area with the parish locations circled (the city of Saskatoon near the bottom of the map includes four parishes).

FIGURE 1
MAP OF PARISH LOCATIONS

The 11 pastors in this diocese are required to make periodic visits to units in order to meet with parish members, provide religious instruction, speak at church services, and train parish leaders. Table 1 lists the first names of each pastor and indicates those who reside outside of Saskatoon (the notion of stewardship parishes will be discussed in the model section).

TABLE 1
LIST OF PASTORS AND STEWARDSHIP PARISHES

Pastor (= resides outside of Saskatoon)**	**Stewardship Parish**
Gerry **	Saskatoon 1st
Kim	Saskatoon 2nd
Larry **	Saskatoon 3rd
Steve L **	Saskatoon 4th
Alan	Flin Flon
Hans	Melfort
Richard	Prince Albert
Steve W	North Battleford

Pastor (** = resides outside of Saskatoon)	Stewardship Parish
John	Meadow Lake
Stan	Kindersley
Armes	None

In the past, a well-meaning diocese volunteer had manually created each year's speaking schedule. Regrettably, such an approach was fraught with complexities for the scheduler and frustration by both the scheduler and the pastors.

This frustration – combined with tough economic times experienced in late 2008 – fueled the desire to seek a better solution to the assignment problem. We note here that the amount of dollars devoted to annual pastor travel reimbursement costs is not strikingly large; indeed, it is measured in the thousands of dollars. However, this non-profit organization wanted to be as frugal as possible in the allocation of its resources. Any budget funds saved would be beneficial to the rest of the religious and charitable work in the diocese. Through a personal contact with one of the pastors, we were approached about providing an analytical solution to this assignment problem.

2.0 METHOD

The optimization model to assign monthly visits of pastors to respective parishes is of the following form:

$$MIN\ 0.42 \sum_{i=1}^{11} \sum_{j=1}^{10} \sum_{k=1}^{12} D_{ij} X_{ijk}$$

subject to:

(1) $$\sum_{j} X_{ijk} \leq 1 \ \forall\, i, k$$

(2) $$\sum_{j} \sum_{k} X_{ijk} \leq 8 \ \forall\, i$$

(3) $$\sum_{j} \sum_{k} X_{ijk} \geq 3 \ \forall\, i$$

(4) $$\sum_{i} X_{ijk} = 1 \ \forall\, j \in J_1;\ k \in Quarter$$

(5) $$\sum_{i} X_{ijk} = 1 \ \forall\, j \in J_2;\ k \in 2\ mth\ pair$$

(6) $$\sum_{i} \sum_{k} X_{ijk} = 10 \ \forall\, j \in J_3$$

(7) $$\sum_{j} X_{ijk} = 0 \ \forall\ i,k \text{ holiday combinations}$$

(8) $$\sum_{k} X_{ijk} = 1 \ \forall\ i,j \text{ stewardship combinations}$$

(9) $$\sum_{k} X_{ijk} \leq 1 \ \forall\ i,j$$

$X_{ijk} \in \{0,1\}$

In a typical fashion to other types of assignment models, we use a "three-subscript" decision variable to represent the visits. Specifically, we have:

X_{ijk} = 1 if pastor i visits parish j in month k
X_{ijk} = 0 otherwise

Pastors are re-imbursed for their (round-trip) travel at a rate of $0.42 per kilometer. Letting D_{ij} represent the round-trip distance between pastor i's home community and the location of parish j, our objective function calculates the total yearly travel cost for this particular diocese.

After speaking with diocese authorities, we learned of the various policies and regulations governing the assignment of pastors to parishes. We note that some of these rules were violated in the previous years' manual schedules; however, church authorities maintained that they desired a set of assignments that observed the constraints depicted in our optimization model. Constraint (1) restricts each pastor to at most one parish visit per month. Constraints (2) and (3) provide lower and upper bounds, respectively, on the annual number of visits per pastor. These constraints were included so as to ensure some degree of equity in annual assignments. Had this constraint not been included, then a cost-minimizing solution would provide those pastors residing in more remote communities almost no monthly assignments.

For constraints (4) through (6), we recognize that specific parishes within this diocese require a certain annual number of pastor visits. Three units in Saskatoon (1^{st}, 2^{nd} and 3^{rd}) need one visit per quarter. For ease of notation, we labeled these three units as J_1. Constraint (4) ensures that each of these parishes is visited quarterly. The Saskatoon 4^{th}, Meadow Lake, Kindersley, Prince Albert and Flin Flon parishes each require one visit every other month. Identifying these units as J_2, we use constraint (5) to assign a sufficient number of pastors to these parishes. Diocese authorities wanted the Melfort and North Battleford parishes (represented by J_3) to receive 10 visits annually. Constraint (6) permits this number of assignments.

Recall that pastors are lay ministers who hold regular occupations during the week. As such, pastors may seek holiday months during particular portions of the year when they would be unavailable for travel assignments. We use constraint (7) to guarantee that pastors are not assigned to visit a parish in accordance with specific holiday requests.

As indicated in Table 1, nearly all of the pastors have specific parishes over which they have particular stewardships. These stewardships were instituted to maintain a connection between the given pastors and parishes. Should the parish require in-depth training or instruction on religious topics, the assigned pastor would be responsible for such work. Constraint (8) guarantees that pastors visit their stewardship units once annually.

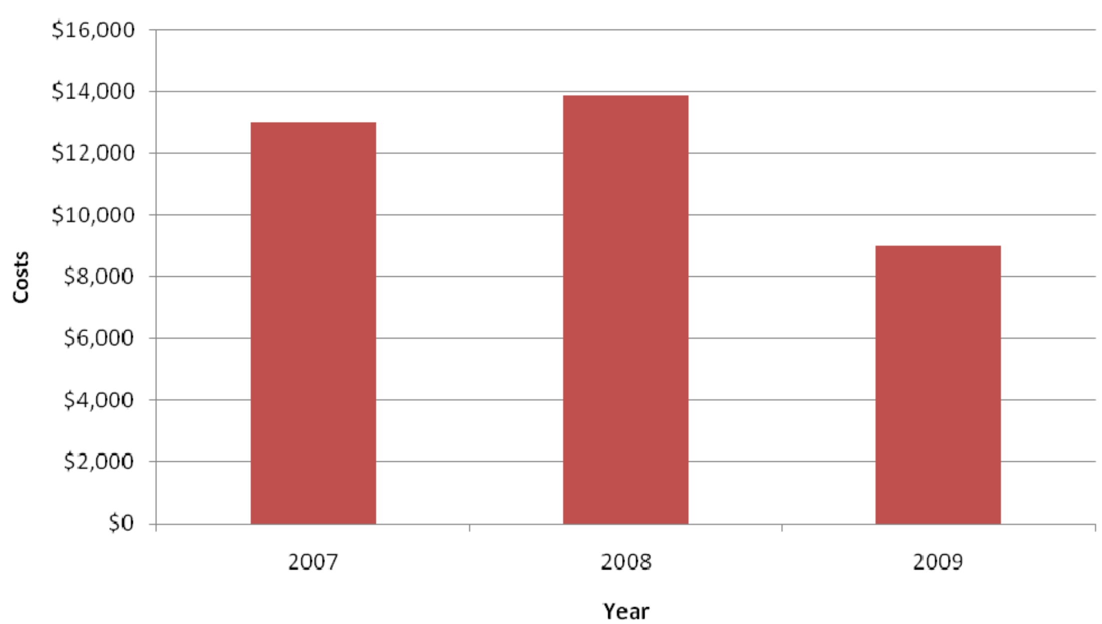
Finally, to obviate the dilemma of assigning a given pastor to the same parish multiple times in year, we use constraint (9) to assure that each pastor can visit a particular parish at most once annually.

3.0 RESULTS

We formulated and solved our integer programming optimization model using *Premium Solver* on an IBM desktop computer with an Intel Pentium D processor. This particular package works within a spreadsheet environment. Our success with such software further demonstrates the applicability of spreadsheet models in exploring real-world analytical problems [1].

The optimizer identified a least-cost set of monthly assignments in a few seconds. The solution it proposed for the 2009 travel schedule reduced costs from those experienced with the manual schedules. For example, in 2007 this diocese incurred a travel reimbursement cost of $13,016. In 2008, the costs climbed to $13,886. The use of our model reduced the 2009 costs to $8,989, a reduction of 35.3% in reimbursed expenses. A graphical comparison of the annual costs is offered in Figure 2.

Figure 2: Annual travel reimbursement costs



The diocese was quite pleased with these results and adopted the proposed assignment schedule in 2009. It is planning to use the results from this optimization model annually in order to determine the most cost-effective set of pastor visits to respective parishes.

4.0 CONCLUDING REMARKS

Our approach has eliminated the hassle and trouble experienced by the scheduler in manually crafting each year's set of assignments. No longer are the parties frustrated by the schedule, and there is a greater sense of equity and fairness in the final product.

Given that the pastors hold regular careers during the week, it is not uncommon for a pastor to relocate outside of the diocese during a calendar year. This necessitates identifying a new pastor to take the relocated pastor's place. Should this happen, one could "lock in" those visits that have already transpired and re-run the optimization model to determine the best set of assignments for the remainder of the year. This provides the diocese with the opportunity to continue to be as cost-effective as possible, even if new pastors are appointed during a year.

REFERENCES

[1] Albright, S. C., Winston, W. L. and Zappe, C. J. *Data Analysis & Decision Making with Microsoft Excel*. Pacific Grove, CA: Thomson Brooks/Cole, 2003.

[2] Bazargan-Lari, M. "Flexible versus fixed timetabling: A case study." *Journal of the Operational Research Society*, 2004, 55, 123-131.

[3] Ernst, A. T., Mills, R. G. J. and Welgama, P. "Scheduling appointments at trade events for the Australian Tourist Commission." *Interfaces*, 2003, 33, 12-23.

[4] Mongell, S. and Roth, A. E. "Sorority rush as a two-sided matching mechanism." *The American Economic Review*, 1991, 81, 441-464.

[5] Roth, A. E. "On the allocation of residents to rural hospitals: A general property of two-sided matching markets." *Econometrica*, 1986, 54, 425-427.

STUDENT SUCCESS ORIENTED EXAMINATION TIMETABLING MODEL

Can Berk KALAYCI, Northeastern University, Boston, MA 02115, (617) 669-8166, canberk@coe.neu.edu
Askiner GUNGOR, Pamukkale University, Denizli, TURKEY, +90. (258) 296-3141, askiner@pau.edu.tr

ABSTRACT

Examination timetabling is a difficult and lengthy problem for which universities devote a large amount of human and material resources every year. In some institutions, administrators generally authorize the final timetable if every exam is scheduled to a time period without any conflicts. Since it is widely accepted that exams are a part of learning process, scheduling exams should also focus on the efficient and healthy way of measuring students' performance. Therefore, it is required to pay attention to many other components during timetabling in order to reach a 'good' schedule. In this study, we specifically focus on students' success when scheduling exams in addition to hard constraints required for feasibility. Student success is positively correlated to the adequate preparation and resting time among exams. Therefore, the main objective of this study is to maximize paper spread (i.e., time gap among exams) considering the difficulties of exams for better success rates of students.

1. INTRODUCTION

A scheduling problem is simply to assign a set of tasks to a set of sources under a set of constraints. In an examination timetabling problem, exams represent tasks and time slots represent sources. Constraints are examined in two categories as hard and soft constraints. Hard constraints have to be assured in order to obtain feasible solutions. It is mostly impossible to satisfy all of the soft constraints. However, the more soft constraints are satisfied the better solution quality for the problem is generally obtained. So it is required to choose soft constraints wisely according to the goal. In this study, we specifically focus on students' success when scheduling exams in addition to several hard constraints required for feasibility, because student success is positively related to the adequate preparation and resting time among exams. One of the desirable attributes of real-life examination timetabling solutions is the maximization of paper spread [1], which is a measure of the amount of study time that each student has among examinations. The amount of study time that students need among examinations is directly proportional to the difficulty of exams. In this study, we focused on the thought that difficult exams demand more study and resting time than the easier ones. Therefore, it is also required to consider the difficulties of exams during the construction of paper spread. When the final schedule is obtained by satisfying this soft constraint as much as possible, students do have enough study and resting time directly proportional to the difficulties of exams. Thus, students' success is maximized in the examination time period.

In the literature, there has been a wide investigation of timetabling approaches in the domain of examination timetabling with various methodologies in order to produce better and feasible timetables. The methodologies are mainly based on artificial intelligence and operational research techniques. Abdullah et al. [2] and Asmuni et al. [3] provide an up to date overview of the research in this field. Carter et al. [4] defined a general equation, presented in Equation 1 that is widely used in examination timetabling research to measure the timetable quality. This objective function, also called proximity cost function, minimizes the penalty for examinations in close proximity trying to spread out each student's schedule. Abdullah et al [2], Burke et al. [5], Petrovic et al. [6], Asmuni et al. [3], Pillay and Banzhaf [7] and Qu et al. [8] have all considered the objective function and its constraints according to this equation.

$$\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{(W_k \times S_{ij})}{S}, \qquad k \in \{0,1,2,3,4\} \tag{1}$$

Where S_{ij} is the number of students involved in both exams E_i and E_j, if $k = |t_i - t_j| < 5$; $W_k = 2^{|4-k|}$ is the cost of assigning two conflicted exams E_i and E_j with k timeslots apart, if $k = |t_i - t_j| < 5$, i.e. $W_1 = 16, W_2 = 8, W_3 = 4, W_4 = 2, W_5 = 1$; t_i and t_j as the timeslots of E_i and E_j, respectively; N is the number of exams and S is the number of students in the problem.

MirHassani [1] used a conflict formulization similar to the this objective function that has three types of exam clashing such as two exams in the same time slot, two exams on the same day and two exams on two consecutive days. We have extended this conflict formulization considering Carter's benchmark model. Naji Azimi [9] has used a high penalty constraint approach which we have also adapted to our study for hard constraints.

In this study, we have specifically extended the general incapacitated examination timetabling problem formulation considering the difficulties of exams in order to modify paper spread in a more efficient way focusing on students' success. The idea of paper spread, which is the amount of studying and resting time that students have among exams, has been extended to the next level with the major effect of exam difficulties.

2. STUDENT SUCCESS ORIENTED FORMULATION

Generally accepted hard constraints for examination timetabling problems can be written as follows:

- No student can sit in more than one exam at the same time.
- All of the planned exams must be scheduled among the available time slots.

So, it is required to assign all exams of each student to different time slots in the available time period. It is required to choose soft constraints carefully according to a defined objective since it is mostly impossible to satisfy all of the soft constraints. In order to increase the quality of a feasible timetable, a critical soft constraint shall be considered here is as follows [1]:

- The maximum amount of time between exams as much as possible among available time slots is necessary.

By the help of the paper spread, students do have more time to study or take a rest between exams. In this study, we focused on the thought that difficult exams require more study and resting time than easier exams. Therefore it is also required to consider the difficulties of exams during the construction of paper spread. Therefore, we modify the soft constraint above as follows:

- Difficult exams should be assigned to considerably far time slots than easier exams according to each other's position.

When the final schedule is obtained by satisfying this soft constraint as much as possible, students do have enough study and resting time directly proportional to the difficulties of exams. Thus, students' success is maximized in the examination time period.

Variables used in the definition and formulation of the examination timetabling problem are given as follows: N is the number of exams, E_i is an exam where $i \in \{1, \dots, N\}$, D is the number of days, T is the given number of available timeslots, t_k specifies the assigned time slot for exam k where $k \in \{1, \dots, N\}$ and$(1 \leq t_k \leq T)$, dc_k specifies the difficulty coefficient for exam k where $k \in \{1, \dots, N\}$ and $(1 \leq dc_k \leq 10)$, 10 represents the most difficult exam and 1 represents the easiest one.

In this problem, we formulate an objective function which tries to space out students' exams throughout exam period by considering the difficulties of each exam. Examination timetabling problem can then be formulated as the minimization of:

In this problem, we formulate an objective function which tries to space out students' exams throughout exam period by considering the difficulties of each exam. Examination timetabling problem can then be formulated as the minimization of:

$$\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (dc_i + dc_j) \times \lambda(t_i, t_j) \tag{2}$$

Where

$$\lambda(t_i, t_j) = \begin{cases} 1000, & if\ t_i = t_j \\ 2^{4-|(t_i - t_j)/D|}, & if\ Mod(t_i, D) = Mod(t_j, D) \\ 1, & if\ |Mod(t_i, D) - Mod(t_j, D)| = 1 \\ 0, & otherwise \end{cases} \tag{3}$$

Subject to:

$$\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} C(t_i, t_j) = 0 \tag{4}$$

Where

$$C(t_i, t_j) = \begin{cases} 1 & if\ t_i = t_j \\ 0 & otherwise \end{cases} \tag{5}$$

Equation (3) presents a proximity value between two exams. For example, if it is accepted that there is three possible time slots for each exam in one day, then the proximity values associated with each time slot may take the values of 1000, 8, 4, 2, 1 or 0. In this equation, the cost for the hard constraint is fixed to 1000 [9]. Equation (4) presents the hard constraint that there are no direct conflicts between exams.

3. A GENETIC ALGORITHM MODEL AS A SOLUTION APPROACH

A genetic algorithm is used to automatically optimize paper spread in this study on the designed mathematical model. Pseudo code for general flow of the proposed algorithm is given in Table 1. Each individual which represents a candidate solution in the population is repeatedly evaluated by genetic operators such as crossover and mutation. In this evaluation process, survival of each individual is determined according to a fitness function. Bad individuals that represent low qualified solutions are destroyed through generations. When the termination criteria are satisfied, the chromosome that has the best gene combination at the last generation represents the best solution found.

Table 1 Pseudo code for general flow of the proposed algorithm

```
SET initial population size to double of population size
SET number of offsprings to crossover rate times population size
SET elit size to elitism rate times population size
CALL generate initial population with initial population size RETURNING initial population
SET number of mutants to chromosome length times mutation rate times population size
FOR each choromosome in the initial population
        CALL determine fitness value with input variables RETURNING fitness value and conflict matrix
ENDFOR
SORT fitness values of each choromosome in initial population in descending order.
DETERMINE general population from initial population
SET trial to zero.
REPEAT
        CALL create mating pool with standart devision coefficient RETURNING Pool
        IF Pool has not enough choromosomes THEN
            RETURN
        ELSE
            CONTINUE
        ENDIF
        CALL crosswhom with Pool and number of offsprings RETURNING MatingSet
        CALL crossover with Population, MatingSet RETURNING Offsprings
        CASE mutation tactic OF
            1: CALL mutate 1 gene with Population, number of mutants, max time slot RETURNING Mutants
            2: CALL mutate 2 genes with Population, number of mutants, max time slot RETURNING Mutants
            4: CALL mutate 4 genes with Population, number of mutants, max time slot RETURNING Mutants
        ENDCASE
        CALL select best of chromosomes with Population, Offsprings, Mutants, Elits RETURNING Population
UNTIL trial < number of generation
```

4. EXPERIMENTAL RESULTS

Simulation results showed that automated approach provides far better solutions than manually constructed timetables according to the desired objective function. Detailed experiment results for five different trials of best experiment sets are shown in Table 2. FF_b : Best fitness value in the first generation, LF_b : Last fitness value in the last generation, t_{cpu} : CPU time occupied during the algorithm's operation and G represents the number of generations in these results of model. A conflict matrix is used to calculate and explain the meaning of each fitness value. We considered five different types of exam clashings for students. In this representation, CT_1: two exams are offered in the same time slot, CT_2 : two exams are offered on two consecutive time slots on the same day, CT_3 : two exams are offered with one time slot break on the same day. CT_4 : two exams are offered with two time slots break on the same day, CT_5 : two exams are offered on two consecutive days.

As the main focus of this study is to maximize students' success, conflict types of CT_5, CT_4 , CT_3, CT_2, CT_1 should be preferred respecively. CT_1 represents the hard constraint satisfaction while the others do soft constraints'. In order to contrast these results, fitness value and conflict matrix of the timetable which is manually constructed is given in Table 3. It can be easily stated that GA model provides far better timetables than manual construction.

Table 2 Detailed experiment results for GA model

FF_b	LF_b	t_{cpu}	G	CT_1	CT_2	CT_3	CT_4	CT_5
6136	2202	2127	118	0	0	0	17	171
6160	2230	2218	123	0	0	1	17	163
6023	2215	2758	153	0	0	2	20	155
5910	2236	2744	152	0	0	0	17	169
6053	2201	2797	155	0	0	1	19	161

Table 3 Fitness value and conflict matrix of manually constructed timetable

LF_b	CT_1	CT_2	CT_3	CT_4	CT_5
9100	0	35	39	7	345

5. CONCLUSION

In this study, we have specifically extended the general incapacitated examination timetabling problem formulation considering the difficulties of exams in order to modify paper spread in a more efficient way of focusing on students' success. The idea of paper spread, which is the amount of studying and resting time that students have among exams, has been extended further to include the major effect of exam difficulties. The focus of student success corresponds that students need maximum amount of resting and studying time among exams depending on the difficulty of exams in the examination period. A genetic algorithm based model was developed and experimental analysis was carried out. Results have proven that considering difficulties of exams in timetabling may increase the students' success. If the students find enough study and resting time among difficult exams, measurement of the teaching outputs may generate objective findings about the students' learning thus leading program improvements. Another benefit of the study is the generation of feasible and efficient timetable quickly using genetic algorithm based model. It minimizes the requirement of the amount of human and material resources at universities each academic year for timetabling.

REFERENCES

[1] Mirhassani, S. A., “Improving paper spread in examination timetables using integer programming.” *Applied Mathematics and Computation*, 2006, 179 (2), 702-706.

[2] Abdullah, S., Ahmadi, S., Burke, E. and Dror, M., “Investigating Ahuja–Orlin’s large neighbourhood search approach for examination timetabling”. *OR Spectrum*, 2007, 29 (2), 351-372.

[3] Asmuni, H., Burke, E. K., Garibaldi, J. M., Mccollum, B. and Parkes, A. J., “An investigation of fuzzy multiple heuristic orderings in the construction of university examination timetables.” *Computers & Operations Research*, 2009, 36 (4), 981-1001.

[4] Carter, M. W., Laporte, G., Lee, S.Y., “Examination timetabling: Algorithmic strategies and applications.” *The Journal of the Operational Research Society*, 1996, 47 (3), 373-383.

[5] Burke, E. K., Mccollum, B., Meisels, A., Petrovic, S. and Qu, R.M., “A graph-based hyper-heuristic for educational timetabling problems.” *European Journal of Operational Research*, 2007, 176 (1), 177-192.

[6] Petrovic, S., Yang, Y. and Dror, M., “Case-based selection of initialisation heuristics for metaheuristic examination timetabling.” *Expert Systems with Applications*, 2007, 33 (3), 772-785.

[7] Pillay, N. and Banzhaf, W, “A study of heuristic combinations for hyper-heuristic systems for the uncapacitated examination timetabling problem.” *European Journal of Operational Research*, 2009, 197 (2), 482-491.

[8] Qu, R., Burke, E. K. and Mccollum, B., “Adaptive automated construction of hybrid heuristics for exam timetabling and graph colouring problems.” *European Journal of Operational Research*, 2009, 198 (2), 392-404.

[9] Naji Azimi, Z., “Hybrid heuristics for Examination Timetabling problem.” *Applied Mathematics and Computation*, 2005, 163 (2), 705-733.

THE FLEXIBLE DELIVERY AND PICKUP PROBLEM WITH TIME WINDOWS

Ying-Yen Chen, National Tsing Hua University, 886-3-5742654, d933810@oz.nthu.edu.tw
Hsiao-Fan Wang, National Tsing Hua University, 886-3-5742654, hfwang@ie.nthu.edu.tw

ABSTRACT

The delivery and pickup problems have drawn much attention in the literature in the past two decades. The current three categories of the delivery and pickup problems had shortages and shortcomings. To conquer these shortages and shortcomings, the flexible delivery and pickup problem was proposed in this paper. The proposed problem remained good application properties of the current three problems, including to mix pickup services and delivery services freely and to reduce the accessing time if the delivery and pickup service of a customer are performed simultaneously. The problem was then formulated into a mixed binary integer programming model and one could implement this model to LINGO to get the optimal solution.

Keywords: delivery and pickup problem, bi-directional logistics, vehicle routing problem with backhauls

INTRODUCTION

With global resources rapidly decreasing, the introduction of reverse logistics may significantly reduce the cost of returned merchandise, improve customer satisfaction, and thereby increase enterprise profit. Recently, many enterprises have incorporated reverse logistics into regularly-operated forward logistics to form a closed-loop supply chain. A state of the art survey of reverse and close-supply chains can be found in Ilgin and Gupta [1]. Within such a loop, the logistics between the distribution/collection center and the customers is the most complicated part because it is related to the bi-directional logistics regarding delivery and pickup activities. In the literature, such problems have been referred to as the delivery and pickup problems (DPP).

DPP applications are frequently encountered, for example, in the distribution system of grocery store chains. Each grocery store may have a demand for both delivery (cf. fresh food or soft drinks) and pickup (cf. outdated items or empty bottles). The foundry industry is another example in Dethloff [2]. Collection of used sand and delivery of purified reusable sand at the same customer location are carried out.

Ropke and Pisinger [3] provided a survey on DPP; referring to the service strategies, DPP is divided into three categories consisting of the vehicle routing problem with backhauls (VRPB), the mixed vehicle routing problem with backhauls (MVRPB) and the simultaneous delivery and pickup problem (SDPP). In the first two kinds of problems, it is assumed that customers can be divided into linehaul customers (customers receiving goods) and backhaul customers (customers sending goods). In VRPB, one considers picking up goods after the last delivery has been made; however in MVRPB, we can mix linehaul and backhaul customers freely within a route. In SDPP, it is assumed that customers simultaneously demand goods from—and supply goods to—the depot. The pickup and delivery should be performed simultaneously such that each customer is visited only once by a vehicle. The simultaneous delivery and pickup operation decreases the customer expenses or inconvenience associated with handling vehicles, but may result in longer routes as Ropke and Pisinger [3] mentioned.

By assigning a time window to each customer, by having travel times associated with each pair of locations, and by having service times associated with the customers, VRPB, MVRPB and SDDP are extended to the vehicle routing problem with backhauls and time windows (VRPBTW), the mixed vehicle routing problem with backhauls and time windows (MVRPBTW) and the simultaneous delivery and pickup problem with time windows (SDPPTW). Based on MVRPBTW (Kontoravdis and Bard [4] and Zhong and Cole [5]) and SDPPTW (Wang and Chen [6]), this paper proposed a new model remaining the flexibility of mixing pickup and delivery operations freely yet addressing the time saving of simultaneously performing delivery and pickup. This kind of problems was referred to as flexible delivery and pickup problem with time windows.

PROBLEM DESCRIPTION

The objective of the flexible delivery and pickup problem is to find a set of routes with the following features that minimizes both the number of vehicle and the total travelling distance.

- All customers demand a delivery service or a pickup service or both.
- Each customer can require his/her delivery or pickup demand be serviced within a specific time window.
- Each demand (no matter pickup demand or delivery demand) can't be split and must be serviced by only one vehicle.
- Each vehicle is restricted by capacity constraints. The combined loads associated with delivery and pickup of each vehicle must not exceed the given capacity.
- Each service consumes a service time; however, if the delivery service and the pickup service of the same customer are performed simultaneously, the accessing time is reduced.
- Vehicles start from a distribution center and end at a collection center.
- Both the distribution center and the collection center have their time horizon.
- A vehicle is allowed to arrive at a customer before the relevant time window, but cannot service the customer until the time window opens.

If the distribution center is the same as the collection center, then CC = DC. In this case, the infrastructure of the system can be seen in Fig. 1. Tiny arrows indicate the delivery services. Dot arrows indicate the pickup services. Bold arrows indicate the simultaneous delivery and pickup services. Each customer node should be pointed by either “one tiny arrow and one dot arrow” or “one bold arrow.”

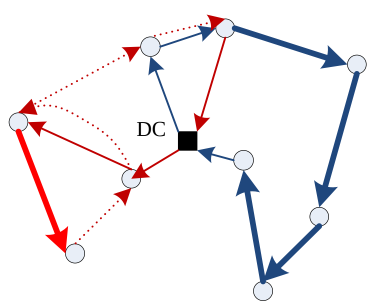


Figure 1: The infrastructure of the flexible delivery and pickup network

MODEL

Assume originally there are n customers. In the proposed model, the original n customers are extended into new $2n$ customers. Each original customer i is divided into two new customers: new customer i and new customer $n+i$. New customer i demands only delivery service and new customer $n+i$ demands only pickup service. The flexible delivery and pickup problem is then formulated into a mixed binary integer programming model denoted by Model FBL (flexible bi-directional logistics) as below.

<u>Model FBL</u>

Notation

Sets

J	Set of all customers, $J = \{ j \mid j = 1,\ldots,2n \}$
J_D	Set of all delivery customers, $J_D = \{ j \mid j = 1,\ldots,n \}$
J_F	Set of all customers plus D.C., $J_F = \{0\} \cup J$
J_R	Set of all customers plus C.C., $J_R = J \cup \{ 2n+1 \}$
J_C	Set of all nodes, $J_C = \{0\} \cup J \cup \{ 2n+1 \}$
V	Set of all vehicles, $V = \{ v \mid v = v_1,\ldots,v_{\lvert V \rvert} \}$

Coefficients

q_v	Capacity of vehicle v, $q_v \in \boldsymbol{R}^+$
g_v	Dispatching cost of vehicle v, $g_v \in \boldsymbol{R}^+$
c_{ij}	Distance between nodes $i \in J_F$, $j \in J_R$; $i \neq j$, $c_{ij} \in \boldsymbol{R}^+$
t_{ij}	Travel time between nodes $i \in J_F$, $j \in J_R$; $i \neq j$, $t_{ij} \in \boldsymbol{R}^+$
d_j	Delivery demand of customer $j \in J$, $d_j \in \boldsymbol{Z}^+$
s_j	Service time of customer $j \in J$, $s_j \in \boldsymbol{R}^+$
r_j	Reduced service time if customer $n+j$ is serviced next to customer j, $j \in J_D$
a_j	Earliest service time of customer $j \in J$, $a_j \in \boldsymbol{R}^+$
b_j	Latest service time of customer $j \in J$, $b_j \in \boldsymbol{R}^+$
a_0	Earliest departure time of any vehicle from D.C., $a_0 \in \boldsymbol{R}^+$
b_{2n+1}	Latest arrival time that a vehicle must return C.C., $b_{2n+1} \in \boldsymbol{R}^+$
M	An arbitrary large constant
α	A parameter indicating the trade-off between dispatching cost and travel cost, $\alpha \in [0,1]$

Decision Variables

L_{0v}	Load of vehicle $v \in V$ when leaving D.C., $L_{0v} \in \boldsymbol{Z}^+$
L_j	Remaining load of a vehicle after having served customer $j \in J$, $L_j \in \boldsymbol{Z}^+$
x_{ijv}	Traveling variable of a vehicle $v \in V$, $x_{ijv} \in \{0,1\}$; if it travels directly from node $i \in J_F$ to node $j \in J_R$, x_{ijv} = 1; otherwise $x_{ijv} = 0$
T_j	Time to begin servicing customer $j \in J$, $T_j \in \boldsymbol{R}^+$
T_{0v}	Departure time of vehicle $v \in V$ at D.C., $T_{0v} \in \boldsymbol{R}^+$

$T_{(2n+1)v}$ Arrival time of vehicle $v \in V$ at C.C., $T_{v(2n+1)} \in \boldsymbol{R}^+$

Model

$$\text{Minimize } z = \alpha \sum_{v\in V}\sum_{j\in J} g_v x_{0jv} + (1-\alpha)\sum_{i\in J_F}\sum_{j\in J_R}\sum_{v\in V} c_{ij}x_{ijv} \tag{1}$$

(Minimize total dispatching cost and total traveling cost)

subject to

$$\sum_{i\in J_F}\sum_{v\in V} x_{ijv} = 1 \qquad \forall j \in J \tag{2}$$

(Service all customer nodes exactly once)

$$\sum_{i\in J_F} x_{ihv} = \sum_{j\in J_R} x_{hjv} \qquad \forall h \in J, \forall v \in V \tag{3}$$

(Arrive at and leave each customer with the same vehicle)

$$\sum_{j\in J} x_{0jv} = \sum_{i\in J} x_{i(2n+1)v} \qquad \forall v \in V \tag{4}$$

(Vehicles which depart from D.C. should finally return C.C.)

$$L_{0v} = \sum_{i\in J_F}\sum_{j\in J} d_j x_{ijv} \qquad \forall v \in V \tag{5}$$

(Initial vehicle loads)

$$L_j \ge L_{0v} - d_j + p_j - M(1 - x_{0jv}) \qquad \forall j \in J, \forall v \in V \tag{6}$$

(Vehicle loads after first customer)

$$L_j \ge L_i - d_j + p_j - M(1 - \sum_{v\in V} x_{ijv}) \qquad \forall i \in J, \forall j \in J \tag{7}$$

(Vehicle loads 'en route')

$$L_{0v} \le q_v \qquad \forall v \in V \tag{8}$$

$$L_j \le q_v x_{ijv} + M(1 - x_{ijv}) \qquad \forall j \in J, \forall v \in V \tag{9}$$

(Vehicle capacity constraints)

$$T_{0v} + t_{0j} - M(1 - x_{0jv}) \le T_j \qquad \forall j \in J, \forall v \in V \tag{10}$$

$$T_i + s_i + t_{ij} - M(1 - \sum_{v\in V} x_{ijv}) \le T_j \qquad \forall i \in J, \forall j \in J - \{n+i\} \tag{11}$$

$$T_i + s_i - r_i + t_{i,n+i} - M(1 - \sum_{v\in V} x_{i,n+i,v}) \le T_{n+i} \qquad \forall i \in J \tag{12}$$

$$T_i + s_i + t_{i(2n+1)} - M(1 - x_{i(2n+1)v}) \le T_{(2n+1)v} \qquad \forall i \in J, \forall v \in V \tag{13}$$

$$a_0 \le T_{0v} \qquad \forall v \in V \tag{14}$$

$$a_j \le T_j \le b_j \qquad \forall j \in J \tag{15}$$

$$T_{(2n+1)v} \le b_{2n+1} \qquad \forall v \in V \tag{16}$$

(Ensure feasibility of the time schedule)

$$x_{ijv} \in \{0,\ 1\} \qquad \forall i \in J_F, \forall j \in J_R, \forall v \in V \tag{17}$$

If the problem scale is not very large, one can solve the problem by implementing this model to LINGO or other software package to get the optimal solution.

DISCUSSIONS

In SDPPTW, it was not allowed to mix pickup services and delivery services freely. In MVRPBTW, if the delivery and pickup service of a customer are performed simultaneously, the accessing time was not reduced. These shortcomings or shortages are conquered in FDPPTW. Furthermore, corresponding to delivery service and pickup service, each customer can request two time windows (delivery time window and pickup time window) in FDPPTW. It is possible that customers request delivering in the morning and picking up in the evening.

It is known that the vehicle routing problem with time windows (VRPTW) is NP-hard (Solomon [7]) and VRPTW is polynomial time reducible from FDPPTW by setting all pickup demands equal to zero. Hence the FDPPTW is also NP-hard and an efficient and effective solution procedure is needed. In Wang and Chen [6], it was noticed that good routes could be inherited from parents to offspring and economical routes would be evolved in a well developed genetic algorithm. We may try to develop a genetic algorithm to solve this problem in the future. Moreover, to demonstrate the computational efficiency and effectiveness of the method, we may implement the method to a bunch of test problems, which was formed by revising Solomon [7] VRPTW benchmarks.

CONCLUSIONS

As the reverse logistics and the closed-loop supply chain networks have been adopted by enterprises, the delivery and pickup problems have also drawn much attention in the literature in the past two decades. The past three categories of the delivery and pickup problems had shortages and shortcomings. It wasn't allowed to mix pickup services and delivery services freely in the vehicle routing problem with backhauls and the simultaneous delivery and pickup problem. On the other hand, if the delivery and pickup service of a customer are performed simultaneously, the accessing time was not reduced in the mixed vehicle routing problem with backhauls. In this study, the flexible delivery and pickup problem was proposed and conquered the above shortages and shortcomings.

The problem was then formulated into a mixed binary integer programming model and one could implement this model to LINGO to get the optimal solution if the scale of the problem is not large. Due to the NP nature of the problem, this study suggested developing a genetic algorithm to get near optimal solutions because it has good mechanism to evolve economical routes. This study also suggested revising Solomon benchmarks to form suitable test problems to test the computational efficiency and effectiveness of methods.

Nowadays, pickup demand often accompanies uncertainty; the fuzzy mathematics may be applied to enhance the model and may provide a broader perspective to the research of the delivery and pickup problem. Hence, determining how to extend the model to cope with uncertainty is a direction of the further study.

ACKNOWLEDGMENT

The authors acknowledge the financial support from the National Science Council with Project no. NSC97-2221-E007-092.

REFERENCES

[1] Ilgin, M.A., Gupta, S.M. "Environmentally conscious manufacturing and product recovery (ECMPRO): A review of the state of the art," *Journal of Environmental Management,* 2010, Vol.91, 563–591.

[2] Dethloff, J. "Vehicle routing and reverse logistics: the vehicle routing problem with simultaneous delivery and pick-up," *OR Spectrum,* 2001, Vol.23, 79–96.

[3] Ropke, S., Pisinger, D. "A unified heuristic for a large class of vehicle routing problems with backhauls," *European Journal of Operational Research,* 2006, Vol.171, 750–775.

[4] Kontoravdis, G., Bard, J.F. "A GRASP for the vehicle routing problem with time windows," *ORSA Journal on Computing,* 1995, Vol.7, 10–23.

[5] Zhong, Y., Cole, M.H. "A vehicle routing problem with backhauls and time windows: a guided local search solution," *Transportation Research Part E,* 2005, Vol.41, 131–144.

[6] Wang, H.F., Chen, Y.Y. "A genetic algorithm for the simultaneous delivery and pickup problems with time window," submitted to *Computer and Industrial Engineering,* 2010.

[7] Solomon, M.M. "Algorithms for the vehicle routing and scheduling problems with time window constraints," *Operations Research,* 1987, Vol.35(2), 254–265.

BENCHMARKING LARGE U.S. RETAILERS USING A DATA ENVELOPMENT ANALYSIS MODEL

Rashmi Malhotra, St. Joseph's University, (610) 660-3497, rmalhotr@sju.edu
D.K. Malhotra, Philadelphia University, (215) 951-2813, MalhotraD@philau.edu
C. Andrew Lafond, College of New Jersey, (215) 951-2950, Lafond@tcnj.edu

ABSTRACT

In this paper, we illustrate the use of data envelopment analysis, an operations research technique, to analyze the financial performance of the seven largest retailers in the U.S. by benchmarking a set of financial ratios of a firm against its peers. Data envelopment analysis clearly brings out the firms that are operating more efficiently in comparison to other firms in the industry, and points out the areas in which poorly performing firms need to improve.

INTRODUCTION

As the financial crisis unfolded in late 2007 and early 2008, most industries started feeling the effects of decreased business and consumer spending. U.S. retail industry is figuring prominently in the global economic meltdown and is currently plagued by bankruptcies, store closures, and layoffs. While retailers that sell "big ticket" items such as cars and furniture are being hit hard, large department stores such as J.C. Penney are reporting as much as a 12% decline in sales from a year ago, September 2007 (Bustillo and Zimmerman, 2008). With this financial crisis, "discretionary spending is drying up as Americans grapple with higher food and gas prices, depressed home values, and diminished retirement accounts"(Bustillo and Zimmerman, 2008). Consumer spending is shifting towards discounters such as Wal-Mart, Target, and Costco however even these discount retailers are feeling the pinch as sales continue to grow at these discount retail stores but at a rate less than expected by analysts. To top it all off, the National Retail Federation forecasts a challenging holiday season with sales gains of 2.2%, well below the ten-year average of 4.4%. Since this crisis began in late summer 2008, the Dow Jones Industrial average has plummeted 20%, the S&P Index is down 15%, and these discount retailers have experienced stock price declines. Wal-Mart has experienced a 12% decline, Costco a 17% decline, and Target a 29% decline in their stock price. All of this is a sign of the current tumultuous times.

During these difficult financial times investors and creditors need insight into a company or industry in order to make a good investment decision. Financial analysis is an excellent tool to help investors make informed investment decisions. Financial statement analysis also helps managers direct the firm as well as help them identify strengths and weaknesses so that management can take advantage of their strengths and minimize the effect of their weaknesses. In this paper we analyze the financial statements of the seven largest discount retailers using the operations management technique data envelopment analysis (DEA). Specifically, we benchmark Wal-Mart, Target, J.C. Penney, Costco, Macys, Sears, and BJ Wholesale against one another using the data envelopment analysis methodology. Through the use of DEA we are able to benchmark the financial ratios of each discount retail store against its competitors. Data envelopment analysis identifies the firms which are performing more efficiently as well as those areas where firms need improvement. The rest of the paper is organized as follows; section II we provide a literature review of previous studies on financial statement analysis, section III discusses the data envelopment analysis model, section IV provides an empirical analysis of our results, and section V summarizes and concludes our study.

LITERATURE REVIEW

Previous studies illustrate the use of data envelopment analysis to evaluate the productivity and efficiency in the retail industry. Sellers-Rubio and Mas-Ruiz (2006) examined the efficiency of supermarket chains in the Spanish retailing industry using DEA. The study evaluated the 100 supermarket chains between 1995 and 2001 revealing that there were "high levels of economic inefficiency in the Spanish retailing industry" (Sellers-Rubio and Mas-Ruiz, 2006). The study found that the "underlying causes of the difference between the current performance level of an organization and the best practically possible level are management style, organizational structure, and product quality among others" (Sellers-Rubio and Mas-Ruiz, 2006). Dontho and Yoo (1998) "measured the relative-to-best performance efficiency of retail outlets characterized by multiple inputs and outputs using data collected from retail stores belonging to a restaurant chain". The greatest advantage of using DEA to evaluate performance efficiency is that a retail outlet is compared to the best performing retail outlets, otherwise known as benchmarking. Through benchmarking, retail outlets of a chain store or a franchise system can be comparatively evaluated for performance, ultimately improving the operation of the entire retail store.

Jiang and Talaga (2006) used DEA to explore the relationship between satisfying customers and building a customer base for the e-tailing industry. The study found that "performance scores for developing a customer base vary across product categories" (Jiang and Talaga, 2006), and "performance score is a good parameter for predicting future change on a unique number of visitors and on the competition pattern for a particular e-tailer" (Jiang and Talaga, 2006). This study used DEA to help e-tailers measure the reach efficiency which is the extent to which the e-tailer is attracting visitors within an online environment and page-view efficiency measures the performance of an e-store in generating more page view per visitor" (Jiang and Talaga, 2006). By simultaneously getting more visitors to an e-tailer's web site and having them stay longer is reflected in the overall efficiency of the unit (Jiang and Talaga, 2006).

Barth (2007) used DEA to show that "new-style retail wine stores with features such as tasting rooms, lecture theatres, and demonstration kitchens used to educate and engage customers have better retail efficiency than old-style stores". The DEA used "sales dollars, labour hours, and litres of inventory depletion from a paired-sample of old-style and new-style facilities to determine the retail efficiency of the stores" (Barth, 2007). The results of the study reflected that the new-style stores had higher retail efficiency than the old style stores and reducing the input in the older stores does not increase the retail efficiency of these stores. Although the study shows that the retail efficiency is increased with the new store features, the contribution of each feature towards the overall improvement in retail performance is unknown (Barth, 2007).

Athanassopoulos and Thanassoulis (1995) used DEA to assess the market effieciency of pubs in the UK to aid in planning. Market efficiency is defined as "the extent to which a unit is exploiting the potential within its catchment area for generating revenue (Athanassopoulos and Thanassoulis, 1995). The inputs are environmental variables as well as one uncontrollable internal variable and the output is the revenue generated (Athanassopoulos and Thanassoulis, 1995).

Pilling, et al (1999) used to adjust salesperson performance for territory characteristics in order to give a more individualized and complete picture of salesperson performance. DEA helped adjust for territory inequities in the evaluation process and identified best practices among a group of salespeople as well as included ways to increase the impact of sales-force related expenditures (Pilling, et al, 1999).
Soterious and Stavrinides (2000) "developed a DEA model to help bank branches use their resources in the most efficient way to improve their service quality, a characteristic thought by many to be the key to gaining competitive advantage and customer loyalty". The results of the study concluded that the "DEA model cannot be used alone to assess branch performance since it only considers a single service quality output which may ignore other important bank branch performance measures, however the model does provide guidelines and direction towards service quality outputs" (Soterious and Stavrinides, 2000)
We did not find any papers or research studies that specifically involve the use of DEA in evaluating the financial performance of the retail industry. This study provides a technique to evaluate retail stores' financial performance through financial ratio analysis using data envelopment analysis model.

MODEL

The Data Envelopment Analysis Model1

Data Envelopment Analysis (DEA) (Charnes et al., 1978) model uses linear programming to measure the comparative performance of different organizational units. Further, this generalized optimization technique measures the relative performance of different decision-making entities that have multiple objectives (outputs) and multiple inputs structure. In the DEA terminology, entities/organization units under study are called Decision-Making Units (DMUs). In our study, the DMUs are the seven retailers under analysis. DEA measures the efficiency with which a DMU uses the resources available (inputs) to generate a given set of outputs. The DEA methodology assesses the performance of the DMU using the concept of efficiency or productivity, defined as a ratio of total outputs to total inputs. Further, the DEA model estimates relative efficiency, which is with reference to the best performing DMU or DMUs (in case multiple DMUs are most efficient). The DEA allocates an efficiency score of unity or 100 percent to the most efficient unit. The low-performing DMUs' efficiency can vary between 0 and 100 percent in comparison to the best performance.

DATA AND METHODOLOGY

According to Standard & Poor's industry survey, liquidity, inventory, and profit margin are critical to a retailer's success.2 Therefore, to study the performance of the retail industry (that includes Wal-Mart, Target, Costco, Macys, Sears, J.C. Penney, and BJ Wholesale), we consider seven financial ratios that have been computed on the basis of information contained in the income statement and balance sheet of these firms. The set of ratios that we use to construct the DEA model are: operations (days of sales outstanding/average collection period, inventory turnover, and asset turnover ratios), profitability (operating profit margin, net profit margin, return on equity and return on assets), and financials (quick ratio and total debt/equity ratio). In order to evaluate a firm's financial performance, a financial analyst usually uses these set of ratios.3 We use the financial statement data available on a quarterly basis from July 2007 to July 2008 from Hoovers Online for this study. Current economic meltdown started in

[1] The main sources of the DEA Model description are Ramanathan (2003) and Zhu (2003).
[2] Standard & Poor's Industry Survey, Retailing: General, May 2009.
[3] See Corporate Finance by Ross, Westerfield, and Jaffe, 8th edition, McGraw Hill/Irwin Publishing Company, 2008.

December 2007. Therefore, this time frame allows us to study the financial performance of the retailers before the crisis as well as after onset of economic crisis. Out of these seven ratios, we specify days of sales outstanding and total debt/equity ratio as input, because for a given company the lower these variables are the better the performance of the company. Similarly, higher operating profit margin, net profit margin, return on equity, return on assets, asset turnover, and inventory turnover imply a better-performing company. Thus, we consider these variables as output variables. Finally, the choice of the DEA model is also an important consideration. We should select the appropriate DEA model with options such as input maximizing or output minimizing, multiplier or envelopment, and constant or variable returns to scale. DEA applications that involve inflexible inputs or not fully under control inputs should use output-based formulations. On the contrary, for an application with outputs that are an outcome of managerial goals, input-based DEA formulations are more appropriate. In addition, for an application that emphasizes inputs and outputs, we should use the multiplier version. Similarly, for an application that considers relations among DMUs, envelopment models are more suitable. Furthermore, the characteristics of the application dictate the use of constant or variable returns to scale. If the performance of DMUs depends heavily on the scale of operation, constant returns to scale (CRS) is more applicable, otherwise variable returns to scale is a more appropriate assumption.

In our study, the comparative evaluation among the companies is an important consideration. Therefore, we select the envelopment models for our analysis. In addition, the outputs are an outcome of managerial goals. Therefore, input-based formulation is recommended for our study. The objective of the analysis is to suggest a benchmark for the seven largest retailing firms. Furthermore, to investigate the effect of scale of operations, if any, among the seven companies, we consider both variable returns to scale and constant returns to scale DEA models. Also, the structure of the DEA model (in envelopment form) uses an equation and separate calculation for every input and output. Therefore, all the input and output variables can be used simultaneously and measured in their own units. In this study, we use the Input-Oriented Variables Return to Scale (VRS) to evaluate the efficiency of seven retail companies. Figure 1 illustrates a decision support system using data envelopment analysis. The decision support system uses the DEA methodology to assess the performance of each company. The DEA-based decision support system uses the company attributes – days of sales outstanding (average Collection Period) and total debt/equity ratio as input variables. The system uses the operating profit margin, quick ratio, return on assets, asset turnover, and inventory turnover as output variables to calculate an efficiency score for a firm. This score is a relative value computed by comparing the given firm to a pool of well-performing companies that serve as a benchmark for the company under evaluation. Each firm is evaluated against the existing firms with an identical set of inputs or outputs that is constructed as a combination of performing and non-performing companies. By using the existing good companies as a "role model," DEA not only helps differentiate well performing (efficient companies from poorly performing (inefficient) firms, but also brings out the reasons why a company may be underperforming.

EMPIRICAL ANALYSIS

Each of the retailers is a homogenous unit, and we can apply the DEA methodology to assess the comparative performance of these companies. We analyze and compute the efficiency of these companies using the financial statements. Using the DEA methodology, we compute an efficiency score for the seven companies on a scale of 1 to 100 on the basis of the financial data for the quarters ending in July 2007, October 2007, January 2007, April 2008, and July 2008. As the DEA methodology does not work with negative values, we first transform the variables by adding the highest negative value to each observation. Table 2 illustrates the efficiency scores for seven companies. Further, we also study the peers (model companies) for inefficient companies.

Table 2 shows the relative performance of the retailers benchmarked against each other. Table 2 also shows that five out of seven companies were ranked as efficient based on the financial data for the quarters ending July 2007, October 2007, January 2007, April 2008, and July 2008 and two companies were inefficient companies. Wal-Mart, Target, Costco, J.C. Penney, and BJ Wholesale are 100% efficient. On the other hand, Macys and Sears are inefficient. Figure 1 shows the efficiency frontier graph of the pooled company data. The 100% efficient companies (blue dots) are on the efficiency frontier, where as the inefficient companies (red dots) are inside the efficiency frontier. The DEA Analyzer calculates the level of inefficiency by measuring the distance between the efficiency frontier and the inefficient companies. Therefore, a financial analyst can use this efficiency frontier to assess the relative efficiency of the firm in the industry. The DEA model compares the days sales outstanding (average collection period), total debt/equity ratio, operating profit margin, return on assets, inventory turnover, quick ratio, and asset turnover ratio.

We present the score in percentage value varying between 0% and 100%. We find that the input efficiency of Wal-Mart, Target, Costco, J.C. Penney, and BJ Wholesale is 100%. On the other hand, the input efficiency of the remaining companies is: Macys (74%), and Sears (64%). This means that the observed levels of operating profit margin, quick ratio, return on assets, inventory turnover, and asset turnover ratio for Macys can be achieved with 74% of the current levels of days sales outstanding (average collection period) and total debt to equity ratio. The same rationale applies to Sears. Five 100% efficient companies turned out to be the best practices companies within the pooled database of the Decision Support System. The best practices companies: Wal-Mart, Target, J.C. Penney, Costco, and BJ wholesale are 100% efficient. As Macys and Sears are inefficient; the next step is to identify the efficient peer group or companies whose operating practices can serve as a benchmark to improve the performance of these companies. Three 100% efficient companies turned out to be the best practices companies within the pooled database of the Decision Support System. Table 3 illustrates the peer group for the inefficient companies.

As shown in Table 3, BJ and J.C. Penney serve as a peer for Sears. In addition, Sears is more comparable to BJ Wholesale (weight 93%) and less comparable to the more distant peer, J.C. Penney (weight 7%). Similarly, Macys had Wal-Mart (weight 55%) and BJ (weight 32%) as its nearest peers that it should emulate. J.C. Penney is a distant peer (weight 13%) for Macys. The efficient peer companies have a similar mix of input-output levels to that of the corresponding inefficient company, but at more absolute levels. The efficient companies generally have higher output levels relative to the company in question. The features of efficient peer companies make them very useful as role models that inefficient companies can emulate to improve their performance. Table 4 illustrates the slack values identified in the next stage of the DEA analysis.

The slack variables for 100% efficient companies are zero. Therefore, Wal-Mart, Target, J.C. Penney, Costco, and BJ Wholesale are Pareto-efficient as the DEA model has been unable to identify some feasible production point which can improve on some other input or output level. On the other hand, for Sears, there is further scope for reducing “Days sales outstanding” by 8.23 units and to improve inventory turnover ratio by 1.78 units. Macys can reduce its “Days sales outstanding” by 4.57 units and improve its inventory turnover ratio by 2.14 units.

SUMMARY AND CONCLUSIONS

In this study, we analyze the performance of seven largest retailers in the United States. We illustrate the use of data envelopment analysis, an operations research technique, to analyze the financial performance of the retail industry by benchmarking a set of financial ratios of a firm against its peers. DEA employs relative efficiency, a concept enabling comparison of companies with a pool of known efficient companies. The DEA model compares a firm with the pool of efficient companies by creating an efficiency frontier of good firms—a tolerance boundary created by establishing the efficiency of firms in terms of several sets of financial ratios. Companies lying beyond this boundary can improve one of the input values without worsening the others. We find that Wal-Mart, Target, J.C. Penney, Costco, and BJ Wholesale are 100% efficient. On the other hand, Macys and Sears are inefficient. We also illustrate the areas in which inefficient companies are lacking behind efficient firms.
We also provide an insight into the benefits of DEA methodology in analyzing financial statements of firms. The decision support system stores the company's historical data, competitive firm's data, and other industry specific data, and uses the DEA methodology to analyze a firm's performance. Moreover, DEA modeling does not require prescription of the functional forms between inputs and outputs. DEA uses techniques such as mathematical programming that can handle a large number of variables and constraints. As DEA does not impose a limit on the number of input and output variables to be used in calculating the desired evaluation measures, it's easier for loan officers to deal with complex problems and other considerations they are likely to confront.

TABLES, FIGURES, & REFERENCES

Tables, figures, references, and full paper available upon request from the authors.

LINEAR PHYSICAL PROGRAMMING FOR SOLVING THE MULTI-CRITERIA DISASSEMBLY-TO-ORDER PROBLEM UNDER STOCHASTIC YIELDS, LIMITED SUPPLY, AND QUANTITY DISCOUNT

Amre Z. Massoud, Northeastern University, Boston, MA, 02115, (617) 373-7635, massoud.a@neu.edu
Surendra M. Gupta*, Northeastern University, Boston, MA, 02115 (617) 373-4846, gupta@neu.edu
(*corresponding author)

ABSTRACT

This paper focuses on the disassembly-to-order (DTO) problem. End-of-life (EOL) products are purchased from a number of suppliers in order to be disassembled into individual components to satisfy the demand for specified numbers of components. However, there are a lot of uncertainties that complicates the process. Previous work in the literature solved the DTO problem under multiple uncertainties and a single objective. Other models solved the multi-criteria DTO problem with a single uncertainty. This paper develops a multi-criteria DTO model that takes into consideration multiple system uncertainties and solves it using Linear Physical Programming (LPP). The main objective was to find the best combination of take-back EOL products to be purchased from every supplier that would satisfy the demand and achieve the aspiration levels of multiple goals. An example is considered to illustrate the model approach.

INTRODUCTION

In the past few years, technology has become more accessible to a broader range of customers due to its declining cost. This has fueled the demand for products with higher technologies and has motivated manufacturing companies to develop new products with better technologies. Hence, consumers were enticed to upgrade their old products, which are still in good working condition, to acquire newer ones with higher technology. This phenomenal of upgrading older products has shortened the life of products and led to the premature disposal of those products. According to the US Environmental Protection Agency (EPA), in 2006 the amount of Municipal Solid Waste (MSW) generated in the US was more than 251 million tons which translates into 4.6 pounds of waste per person per day.

Waste generated from the premature disposal of end-of-life (EOL) products has had a huge negative impact on the environment. Raw materials are depleting at high rate, harmful waste is increasing at high rate, and landfills are filling up rapidly. Consequently, it is crucial to properly manage waste from EOL products in order to minimize the negative impact on the environment.

Around the world, new laws and regulations have been established by governments which aim to control and reduce the amount of waste being sent to landfills. In addition to the new laws and regulation, consumers are becoming more aware of the problems and are putting a lot of pressure on manufacturers to be more green. This has led manufacturing companies to become more environmentally conscious and come up with different solutions to deal with this problem. For example, returned EOL products can remanufactured, reused, recycled, or disposed of. Remanufacturing, reusing, and recycling EOL products can help decrease the rate of depletion of virgin nature resources and reduce the amount of harmful waste sent to landfills. Additionally, this will reduce the manufacturer's dependency on the natural resources. However, disposal of EOL products contributes to the problem and therefore should be used as a last resort.

In order for manufacturers to remanufacture, reuse, and recycle EOL products, they first need to estimate the number of EOL products needed and obtain them though the disassembly process. After that, EOL products need are disassembled into individual components and subassemblies in order to satisfy the different demands. However, there a lot of uncertainties that exists in the process. Those uncertainties in the process make it extremely difficult to know in advance the required number of EOL products needed for disassembly. Researchers have developed a number of models and techniques that address the different uncertainties in the process. However, they suffer from some limitation. Most previous models consider only one uncertainty which was the stochastic yield of components from EOL products. However, there are other uncertainties in the process which have not been addressed. For example, EOL products could be purchased from multiple suppliers, suppliers might offer discounts rates on the total purchase to increase their competitiveness, independent outside suppliers might offer ready-to-use components, and manufacturers might have capacity limit on the number of components that can be carried into inventory. A few models have been developed in the literature which considered more than one uncertainty but were solved under a single objective function.

Our paper develops a multi-criteria disassembly-to-order (DTO) model that takes into consideration the different uncertainties mentioned. The DTO model is solved using Linear Physical Programming (LPP). We focus on the DTO system where a variety of EOL products are purchased for disassembly. The model's main objective is to find the best combination of EOL products to be purchased from every supplier which satisfies the demand while achieving the desired levels of multiple goals. In our model, the decision maker (DM) might find it difficult to specify the desired levels, weights, and priority of each goal. As a result, LPP is used to remove the DM from the process of determining the priority of each goal and handles the vagueness of the desired levels in the model.

LITERATURE REVIEW

The area of disassembly has been the interest of many researchers. The recent book by Lambert and Gupta [1] gives a good general understanding of the area of disassembly. Gungor and Gupta [2] presented a comprehensive study of the different issues in environmentally conscious manufacturing and product recovery. Kongar and Gupta [3] presented a multi-criteria decision making approach for DTO system where the objective was to find the best combination of EOL products to be disassembled that would fulfill the demand under a number of financial, environmental, and physical constraints. Imtanavanich and Gupta [4] modeled the DTO problem with stochastic yields using multi-criteria decision making approach by converting the stochastic yields into their deterministic equivalents using heuristic procedures. Imtanavanich and Gupta [5] solved the DTO problem using evolutionary computation with LPP. Massoud and Gupta [6] considered the multi-period DTO problem under stochastic yields, limited supply, and quantity discount with the objective of finding the best combination of EOL products to be taken-back that would maximize total profit. Kongar and Gupta [7] focused on the DTO system for EOL electronic products where electronic products were taken-back for disassembly and proposed a LPP model that considered a number of financial, environmental, and performance goals.

DISASSEMBLY-TO-ORDER SYSTEM

In a DTO system, a wide range of EOL products are taken-back or purchased from the end user in order to satisfy the demand for components and materials. Before the demand can be satisfied, EOL products need to be disassembled into individual components and subassemblies. Two types of disassembly are used to take apart EOL products: complete and selective disassembly. In complete disassembly, all items are disassembled; while in selective disassembly, only items which are of interest are disassembled. Moreover, the process of disassembly can be divided into two types: destructive and non-destructive. Destructive disassembly process allows for a few items to be damaged during the disassembly process, while the non-destructive disassembly doesn't allow for items to be damaged during the disassembly process.

PROBLEM DEFINITION

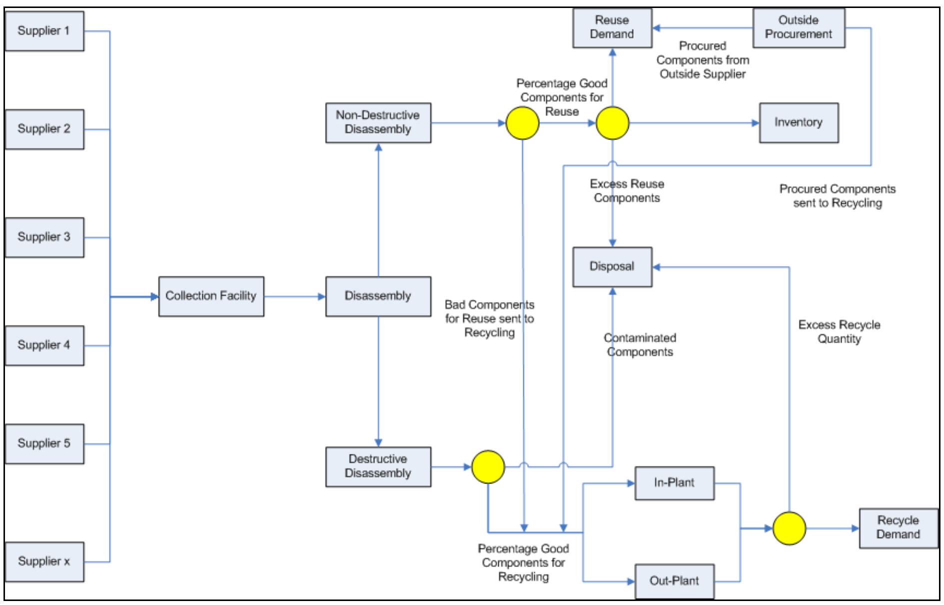


Figure 1: Disassembly-to-Order System

Figure 1 gives a graphical representation of the DTO system. The DTO process starts with the purchased of EOL products from a number of suppliers. The collection facility will then collect the acquired EOL products in order to be inspected and prepared for disassembly. Ready-to-disassemble products are then sent to the disassembly process. The type of disassembly process will depend on the final use and condition of components. For example, reuse demand and stored components will require the use of non-destructive disassembly process. On the other hand, recycling demand will require the use of destructive disassembly process.

Component yield from both non-destructive and destructive disassembly process is stochastic. Components obtained from both the non-destructive and destructive disassembly process are sorted, inspected, and divided into two groups. Components from the non-destructive disassembly process are divided into good and bad non-destructive components based on their functionality. Good non-destructive components are used to satisfy the demand for reused components and storage while bad non-destructive components are sent to the recycling process.

If reuse demand cannot be met from the good non-destructive components, additional components may be procured from an outside supplier at a premium cost since demand shortage is not allowed. Components from the destructive disassembly process are divided into good and bad (contaminated) destructive components. Good components from the destructive process along with bad components from the non-destructive process are sent to the recycling processing facility where they are recycled either in-plant or out-plant. In-plant recycling is costs less but has a capacity limit on the number of components it can process. Therefore, any additional components beyond the limit of the in-plant facility is sent to the out-plant facility. Finally, excess components from all processes are disposed of along with the contaminated components.

This paper develops a multi-criteria DTO model that takes into consideration a number of system uncertainties and solves the model using LPP. The main objective is to find the best combination of EOL products to be purchased from every supplier which satisfies the demand while achieving the desired levels of multiple goals. In this case, the decision maker (DM) might find it difficult to specify the desired levels, weights, and priority of each goal. Therefore, LPP is used to remove the DM from the process of determining the priority of each goal and handles the vagueness of the desired levels in the model.

NUMERICAL EXAMPLE

In order to illustrate the application of the multi-criteria DTO model, we present a numerical example in this section. The example considers 2 suppliers, 6 products, and 10 components. Each supplier offers 3 different types of products where each product consists of 8 different components and the prices and conditions of products differ from one supplier to the other. Discount schedules are offered by each supplier and they differ based on the total dollar amount purchased from each supplier. Tables 1 though 6 summarize the input data for our model. In addition to the data presented in the tables below, the following data is needed to solve the model: in-plant recycling cost = $1.00, out-plant recycling cost = $1.80, contaminated disposal cost = $2.00, and inventory capacity = 50 for each component.

Table 1: Product Purchase Price & Capacity

Product	Supplier 1		Supplier 2	
	Purchase Price	Capacity	Purchase Price	Capacity
1	$20	150	$24	186
2	$21	152	$20	170
3	$22	181	$22	168

Table 2: Component Yields from Both Suppliers

		A	B	C	D	E	F	G	H	I	J
	Produc 1	0.00	0.67	0.62	0.81	0.83	0.54	0.00	0.53	0.54	0.65
Supplier 1	**Produc 2**	0.68	0.00	0.61	0.79	0.82	0.59	0.65	0.27	0.00	0.29
	Produc 3	0.77	0.75	0.54	0.00	0.00	0.55	0.62	0.55	0.52	0.64
	Produc 1	0.00	0.80	0.70	0.94	0.96	0.61	0.00	0.60	0.64	0.74
Supplier 2	**Produc 2**	0.75	0.00	0.71	0.96	0.99	0.68	0.70	0.31	0.00	0.34
	Produc 3	0.84	0.92	0.64	0.00	0.00	0.65	0.66	0.64	0.65	0.75

Table 3: Discount Schedule Supplier 1

Discount Schedules	Discount Rate
0 < 2,500	0.0%
2,500 < 5,000	3.0%
5,000 < 8,000	5.0%
>= 8,000	8.0%

Table 4: Discount Schedule Supplier 2

Discount Schedules	Discount Rate
0 < 2,500	0.0%
2,500 < 5,000	2.5%
5,000 < 8,000	6.0%
>= 8,000	7.5%

Table 5: Component Input Data

Component	Reuse Demand	Reuse Sell	Reuse Perc.	Recycle Demand	Recycle Sell	Recycle Perc.	Holding Cost	Proc. Cost	Non-Dest. Cost	Dest. Cost	Disposal Cost
A	240	$14	80%	170	$8	98%	$2.10	$8	$0.55	$0.35	$1.80
B	240	$15	85%	135	$10	95%	$2.05	$8	$0.50	$0.30	$1.70
C	330	$13	85%	145	$10	99%	$2.00	$7	$0.45	$0.25	$1.90
D	180	$11	80%	135	$8	96%	$2.10	$6	$0.50	$0.30	$1.90
E	185	$13	75%	170	$9	98%	$2.15	$8	$0.60	$0.40	$1.95
F	335	$13	90%	155	$9	97%	$2.20	$7	$0.65	$0.35	$1.75
G	240	$13	85%	145	$10	97%	$2.15	$9	$0.60	$0.30	$2.00
H	330	$12	85%	215	$9	99%	$1.95	$8	$0.55	$0.35	$1.70
I	195	$13	80%	155	$8	99%	$2.00	$7	$0.50	$0.30	$1.80
J	355	$14	80%	215	$9	99%	$2.15	$8	$0.45	$0.25	$1.85

Table 6: Range Limits for each Goal

Objective Function 1: Profit				**Objective Function 2: Procurement Cost**			
Class-2S				**Class-1S**			
Performance Level	Range	Limits	Values	Performance Level	Range	Limits	Values
Ideal	≥ t11-	t11-	29,000	Ideal	≤ t11+	t21+	3,000
Desirable	(t12- , t11-)	t12-	28,500	Desirable	(t12+ , t11+)	t22+	3,250
Tolerable	(t13- , t12-)	t13-	28,000	Tolerable	(t13+ , t12+)	t23+	3,500
Undesirable	(t14- , t13-)	t14-	27,500	Undesirable	(t14+ , t13+)	t24+	3,750
Highly Undesirable	(t15- , t14-)	t15-	27,000	Highly Undesirable	(t15+ , t14+)	t25+	4,000
Unacceptable	≤ t15-			Unacceptable	≥ t15+		
Objective Function 3: Purchase Cost				**Objective Function 4: Disposal Cost**			
Class-1S				**Class-1S**			
Performance Level	Range	Limits	Values	Performance Level	Range	Limits	Values
Ideal	≤ t11+	t31+	12,000	Ideal	≤ t11+	t41+	1,000
Desirable	(t12+ , t11+)	t32+	12,250	Desirable	(t12+ , t11+)	t42+	1,050
Tolerable	(t13+ , t12+)	t33+	12,500	Tolerable	(t13+ , t12+)	t43+	1,100
Undesirable	(t14+ , t13+)	t34+	12,750	Undesirable	(t14+ , t13+)	t44+	1,200
Highly Undesirable	(t15+ , t14+)	t35+	13,000	Highly Undesirable	(t15+ , t14+)	t45+	1,250
Unacceptable	≥ t15+			Unacceptable	≥ t15+		

In order to solve the multi-criteria LPP problem, we used Lingo 11.0. The solved model yields the following total purchase and discount rates for supplier 1 and 2 respectively: $6,603 with 5.0% discount and $5,400 with 6.0% discount. The combination of products 1, 2, and 3 purchased from Supplier 1 were: 104, 31, and 176; while the combination of products 1, 2, and 3 purchased from Supplier 2 were: 132, 104, and 116. Table 7 below summarizes other output data from the model which include: the quantity of procured components for reuse and recycle, quantity of stored components, and quantity of disposed components from reuse, recycle, and contamination.

Table 7: Model Output for LPP problem

Component	Outside Procurement		Ending Inventory	Disposal Quantity		
	Reuse Procurement	Recycle Procurement		Excess Reuse	Excess Recycle	Contaminated Quantity
A	0	11	26	0	1	2
B	2	1	46	0	5	5
C	34	0	0	0	119	2
D	1	56	32	0	0	1
E	1	84	0	0	1	1
F	32	0	0	0	98	7
G	3	0	0	0	41	5
H	87	0	0	0	102	3
I	0	0	0	0	76	2
J	87	0	0	0	78	2

CONCLUSION

The multi-criteria disassembly-to-order (DTO) problem was evaluated in this paper. End-of-life (EOL) products were purchased from a number of suppliers in order to be disassembled into individual components to satisfy the different demand. The main objective was to find the best combination of EOL products to purchase from every supplier that satisfies the demand while achieving the desired levels of multiple goals. Linear Physical Programming (LPP) was used to remove the decision maker (DM) from the process of determining goal priorities and to handle the vagueness in the model.

REFERENCES

[1] Lambert, A. J. D. and Gupta, S. M., "*Disassembly Modeling for Assembly, Maintenance, Reuse, and Recycling*," CRC Press, Boca Raton, Florida, ISBN: 1-57444-334-8, 2005.

[2] Gungor, A. and Gupta, S. M., "Issues in environmentally conscious manufacturing and product recovery: a survey," *Computers & Industrial Engineering*, Vol. 36, No. 4, 811-853, 1999.

[3] Kongar, E. and Gupta, S. M., "A multi-criteria decision making approach for disassembly-to-order systems," *Journal of Electronics Manufacturing*, Vol. 11, No. 2, 171-183, 2002.

[4] Imtanavanich, P. and Gupta, S. M., "Calculating disassembly yields in a multi-criteria decision making environment for a disassembly-to-order system," *Proceedings of the 2005 Northeast Decision Sciences Institute Conference*, Philadelphia, Pennsylvania, 2005.

[5] Imtanavanich, P. and Gupta, S. M., "Evolutionary computation with linear physical programming for solving a disassembly-to-order system," *Proceedings of the SPIE International Conference on Environmentally Conscious Manufacturing VI*, Boston, Massachusetts, 30-41, 2006.

[6] Massoud, A. Z. and Gupta, S. M., "Solving the multi-period disassembly-to-order system under stochastic yields, limited supply, and quantity discount," *Proceedings of 2008 ASME International Mechanical Engineering Congress and Exposition*, Boston, MA, 2008.

[7] Kongar, E. and Gupta, S. M., "Solving the disassembly-to-order problem using linear physical programming," *International Journal of Mathematics in Operational Research*, Vol. 1, No. 4, 504-531, 2009.

BAYESIAN UPDATING OF RECOVERY RATE DISTRIBUTION

H. Kıvanç Aksoy, Eskişehir Osmangazi University,Turkey, 90(222)239-3750, hkaksoy@ogu.edu.tr

ABSTRACT

We consider an inventory system with a single item remanufacturable returns and stochastic demand. The aim of this paper is to use Bayesian methods in dealing with recovery rate uncertainty of the used products in the remanufacturing system. In Bayesian analysis the new information is combined with the previously available information. At this point the prior information (distribution) corresponds to the historical data or the subjective thought of the decision maker about the random parameter of the involved process. The consequential decision or inferential statement (posterior distribution) pooled all available information about the uncertain parameter of the interest. A numerical example is presented to show the effect of the parameter revising on the system's operating cost performance.

Keywords: Bayesian update; Remanufacturing; Performance evaluation; Queueing network; Throughput

INTRODUCTION AND LITERATURE REVIEW

Remanufacturing is a significant constituent of product recovery opportunities. Product recovery management is focused on the collection of used and discarded products and the searching of the opportunities to remanufacture the products/assemblies reuse the components or recycle the materials. The objective of product recovery management, as stated by [13] is "to recover as much of the economic (and ecological) value as reasonably possible, thereby reducing the ultimate quantities of waste". Remanufacturing is one of the most desirable options of product recovery. Remanufacturing operations tend to be labor intensive that lead to significant variability in the processing times and recovery rate of the retrieved products at various shop floor operations. The uncertainties surrounding the returned products further complicate the modeling and analysis of product recovery problems. As such, forecasting the quantity, quality levels of used products, recovery rate of the process and remanufacturing operations lead time are difficult. The results of the stated uncertainties lead to undersupply or obsolescence of systems serviceable inventory, and improper remanufacturing plan.

There is an expanding body of literature in the area of remanufacturing/manufacturing systems and product recovery. Ilgin and Gupta [6] review the literature in the area of environmentally conscious manufacturing and product recovery. The book by Pochampally et al. [11] explores various areas of reverse logistics and closed-loop supply chains. They evaluate the futurity of used product by Bayesian updating from expert systems approach. Krikke *et al.* [9] investigated one of the vital management problems of the original equipment manufactures that is to determine to what extent returned product must be disassembled and which recovery and disposal options should be decided on. Inderfurth *et al.* [7] considered a periodic review model which is based on different remanufacturing options and disposal option for the product recovery in stochastic remanufacturing systems. Aras et al. [1] presented an approach for assessing the impact of quality-based categorization of used products. The authors showed that incorporation of returned product quality in the remanufacturing and disposal decisions can lead to major cost savings for the hybrid systems

The Bayesian revising methodology is computationally straightforward in the existence of completely obtainable return or recovery data. In Bayesian analysis the new information is combined with the previously available information. At this point the prior information (distribution) corresponds to the historical data or the subjective thought of the decision maker about the unknown parameter of the involved process. The consequential decision or inferential statement (posterior distribution) pooled all available information about the uncertain parameter of the interest. The performance of the following updates depends on the prior information therefore the determination of prior information is significant. If there isn't any basis of the prior information then the decision maker may consider the non-informative priors about the random variable which represents the unrevealed parameter may obtain any value in its domain equally likely [4], [5], [12], and [15].

The model that we've investigated is based on expected recovery rate characterization and derived order quantity resulting from this approach. We assume that single production/outside procurement order decision made before the period to persuade the unfulfilled demand at the end of the period is considered as lost sales [10], [14]. The primary objective of this paper is obtaining optimal order quantity and optimal system's expected cost in the presence of stochastic recovery rate distributions with Bayesian approach. Also we have analyzed the effect of recovery rate adjustment on the cost function with various return and demand rate combinations.

REMANUFACTURING MODEL AND ANALYSIS

Throughout the paper, for recovery potential we use the term recovery rate (r_i) to refer to the acceptable fraction of used parts that satisfies the quality specifications for remanufacturing. We assume that the returned products recovery rate probability distribution functions are mutually independent among the stations. Also, it's assumed that recovery rate distributions are independent with the magnitude of the returned rate flow. The p.d.f. and c.d.f. of the re-usable rate at stage i is denoted as $f(r_i)$ and $F(r_i)$ respectively. The demand rate represented by γ with c.d.f $Q(\gamma)$ and p.d.f. $q(\gamma)$.

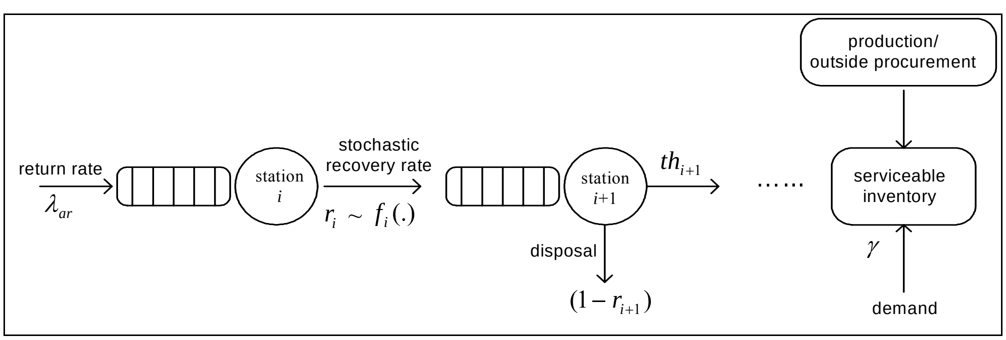


Figure 1. Remanufacturing system with stochastic recovery rate.

The market environment for the remanufactured products is followed the “make-for-stock” policy. In the make-for-stock environment all customer demand is planned to be met from finished serviceable inventory. The detailed mechanism of the model is depicted in Figure 1. We model the remanufacturing system as an open queueing network and use the decomposition principle and expansion methodology to analyze it. To obtain the approximate throughput rates (TH_i) of each server and the entire remanufacturing network, we utilize the expansion methodology. Details of the method and necessary derivations for production lines can be obtained from [2], [3], and [8].

We model the remanufacturing system as an open queueing network and use the decomposition principle and expansion methodology to analyze it. To obtain the approximate throughput rates (TH_i) of each server and the entire remanufacturing network, we utilize the expansion methodology. Details of the method and necessary derivations for production lines can be obtained from [2], [3], and [8].

Bayesian Update

In this section, we present the essentials of the Bayesian updating of the core recovery rate distribution in brief.
A remanufacturing system can be modeled as a collection of various service areas where jobs arrive at different rates and demand services with unequal processing times. The remanufacturing system reclaims a single part product and that item has a fixed or predetermined routing through the system. The goal of the remanufacturing process is to bring the product at the original quality level as manufactured first time. After the remanufacturing operations, items are directed to the serviceable inventory from where the demand is satisfied We assume that the demand rate, γ is greater than the product return rate, λ_{ar}. Thus, outside procurement is needed to supplement any additional demand. It is assumed that when the demand is not satisfied, a lost sales cost is incurred. Similarly, when the demand is less than the inventory level, an inventory holding cost is incurred.

In the remanufacturing process, a low reusable rate of products does not, by itself, create a major problem since it is always possible to supplement the deficient amount with new parts/materials bought from outside suppliers to satisfy the demand in a given period. However, the variable recovery rate of a product complicates production and inventory planning.

In Figure 1, *r* represents the reusable rate of the parts (cores), which are disassembled from returned products. Generally, reusable rate *r* can be defined as $r = 1 - sr$, where *sr* corresponds to the scrap rate. *r* is a stochastic variable with mean $\bar{r}$, standard deviation σ_r, and density function $f(r)$ such that $0 \le r \le 1$. If we assume that the distribution of *r* is symmetrical, then clearly;

$$P(X_{rem}(t) < X_{req}(t)) = P(X_{rem}(t) > X_{req}(t)) = 0.5$$

where

$X_{rem}(t)$: denotes the output of remanufacturing operations in period *t* and

$X_{req}(t)$: denotes the required amount of remanufactured item to satisfy the demand in period *t*.

If R_p represents the planned reusable rate (expected recovery rate of returned products), then the following probabilities hold:

$P_{und} = P(X_{rem} < X_{req}) = \int_0^{R_p} F(r)dR_p$ (the probability of undersupply of remanufactured parts)

$P_{obs} = P(X_{rem} > X_{req}) = 1 - \int_0^{R_p} F(r)dR_p$ (the probability of obsolescence of remanufactured parts)

where $F(r)$ is the distribution function of r.

We considered that core recovery rate in consecutive periods are independent and identically distributed (i.i.d) stochastic variables depicted by r_i for periods $i = 1, 2, 3$ . . . . In a remanufacturing/manufacturing system recovery rate, r_i, is a decision variable and this uncertain parameter has continuous density function $f(r_i)$. According to Bayesian approach decision maker's preliminary statements regarding to this stochastic variable expressed by a prior distribution. Additionally, the sample statistic y_s provides the major information about the uncertain parameter of interest, r_i. In Bayesian analysis the new information is combined with the previously available information.

If random variable r_i is continuous, the prior and posterior distribution of r_i can be represented by density functions. The posterior distribution is the conditional density of r_i, given the observed value , y, of the sample statistic y_s.
The posterior density can be written as follows [15].

$$f(r_i|y_s = y) = \frac{f(r_i) f(y|r_i)}{\int_{-\infty}^{\infty} f(r_i) f(y|r_i) dr_i} \tag{1}$$

The densities $f(r_i|y_s = y)$ and $f(r_i)$ represent the posterior and prior distributions respectively and $f(y|r_i)$ represents the likelihood function. In the continuous parameter case, Bayes' theorem can be expressed in words as;

$$\text{posterior density} = \frac{(\text{prior density}) \times (\text{likelihood})}{\int (\text{prior density}) \times (\text{likelihood})}$$

It should be noted that in Equation (1), it's assumed that r_i is a continuous random variable whereas the sample statistics y_s may be discrete or continuous.

NUMERICAL ANALYSIS

In this section we examine the effect of recovery rate update on the performance of the system. We assume that single production order decision made before the period to persuade the demand and sold products may be returned by the customer to the remanufacturing system Unfulfilled demand at the end of the period is considered as lost sales. The primary objective of this paper is obtaining optimal order quantity and associated total system cost in the presence of stochastic recovery rate using Bayesian approach. Additionally, we have analyzed the effect of recovery rate modification on the total cost function with various stages of product life cycle. To assess the effect of recovery rate on the remanufacturing system performance we defined the expected cost per period, which comprises of the variable and fixed remanufacturing and outside procurement costs, disassembly, testing, disposal and remanufacturing costs, holding costs for serviceable inventory and lost sales costs. The following notation is used for the steady state cost function.

$$E(TC|r_i) = c_p E(RP) + c_d E(D) + c_t E(T) + c_{dis} E(Dis) + c_h E(Inv) + c_l E(Ls) + c_r E(R) + c_m E(OP)$$

$E(RP)$:	expected number of returned products per unit of time.
$E(D)$:	expected number of disposed items per unit of time.
$E(T)$:	expected number of tested products per unit of time.
$E(Dis)$:	expected number of disassembled products per unit of time.
$E(Inv)$:	expected on hand inventory level per unit of time.
$E(Ls)$:	expected lost sales per unit of time.
$E(R)$:	expected number of remanufactured products .
$E(OP)$:	expected number of new products. .
c_p :	purchase cost of returned product ($/product).
c_d :	disposition cost per product ($/product).
c_t :	testing cost per returned product ($/product)
c_{dis} :	disassembly cost per returned product ($/product)
c_h :	inventory holding cost per returned product ($/product/time)
c_l :	lost sales cost ($/product/time)
c_r :	remanufacturing operations cost at station j ($/product)
c_m	cost of new product ($/product).

This section presents a numerical study on the cost function with Bayesian update for the recovery rate distribution. The cost variables were assumed to be as follows:

$c_m = 25,\ c_p = 4,\ c_{dis} = 6,\ c_d = 5,\ c_r = 8,\ c_h = 1,\ c_l = 5,\ c_t = 1,$

We assumed that recovery rate of the returned products to the system is normally distributed with unknown mean μ_r and known variance $\sigma_r^2 = 0.25$. Furthermore, the prior distribution about the unknown parameter μ_r considered has a normal density with (μ_r' =0.65, $\sigma_r'^2$ =0.2). Based on sample data with y_s the posterior distribution of r_i is found from Bayesian updating with normal prior density as follows;

$$f''(r_i) = \frac{1}{\sqrt{2\pi\sigma_r''^2}} e^{-(r_i-\mu_r'')^2/2\sigma_r''^2} \qquad \forall r_i \tag{2}$$

where,

$$\mu_r'' = (\sigma_r^2\sigma_r'^2)/(\sigma_r^2 + \sigma_r'^2) \text{ and } \sigma_r''^2 = (\mu_r'\sigma_r^2 + y\sigma_r'^2)/(\sigma_r^2 + \sigma_r'^2) \tag{3}$$

In order to reflect the effect of Bayesian approach on the system operation cost, average serviceable inventory and process times different combinations of demand and return rates are utilized and the results are summarized in Table 1.

Table 1. System performance measures with different system parameters

	$\lambda_{ar} = 0.1 \quad \gamma = 0.3$	$\lambda_{ar} = 0.25 \quad \gamma = 0.6$	$\lambda_{ar} = 0.8 \quad \gamma = 1.0$	$\lambda_{ar} = 0.7 \quad \gamma = 0.5$
$E[TC \mid \mu_r]$	29.9046	28.5783	28.0186	29.0106
$E[TC \mid \mu_r'']$	29.3915	28.214	25.7682	27.0953
$PT \mid \mu_r$	5.1625	7.18199	11.6356	10.8695
$PT \mid \mu_r''$	5.18829	7.08454	11.6138	10.553
$SInv. \mid \mu_r$	6.07708	6.312262	7.48509	8.35634
$SInv. \mid \mu_r''$	5.40161	6.3069	7.20807	8.09779

REFERENCES

[1] Aras, N., Boyacı, T., Verter, V., “The effect of categorizing returned products in remanufacturing”, *IIE Transactions*, 2004, 36, 319-331.

[2] Gupta, S. M. and Kavusturucu, A., “Modeling of finite buffer cellular manufacturing systems with unreliable machines”, *International Journal of Industrial Engineering*, 1998, **5**(4), 265-277.

[3] Gupta, S. M. and Kavusturucu, A, “Production systems with interruptions, arbitrary topology and finite buffers”, *Annals of Operations Research*, 2000, 93, 145-176.

[4] Hill, R. M., “Applying Bayesian Methodology with A Uniform Prior to the Single Period Inventory Model”, *European Journal of Operations Research*, 1997, 98, 555-562,

[5] Hill, R. M., “Bayesian Decision Making In Inventory Modeling”, *IMA Journal of Mathematics Applied in Business and Industry*, 1999, 10, 147-163.

[6] Ilgin, M. A. and Gupta, S. M., "Environmentally Conscious Manufacturing and Product Recovery (ECMPRO): A Review of the State of the Art", *Journal of Environmental Management*, 2010, 91(3), 563-591.

[7] Inderfurth, K., de Kok, A. G., Flapper, S. D. P., “Product recovery in stochastic remanufacturing systems with multiple reuses options”, *European Journal of Operations Research*, 2001, 133, 130-152.

[8] Kavusturucu, A. and Gupta, S. M., “Analysis of Manufacturing flow lines with unreliable machines”, *International Journal of Computer Integrated Manufacturing*, 1999, 12(6), 510-524.

[9] Krikke, H. R., Van Harten, A., Schuur, P. C.,. “On a medium term product recovery and disposal strategy for durable assembly products”, *International Journal of Production Research*, 1998, 36(1), 111-139.

[10] Mostard, J., and Teunter, R., “The Newsboy Problem with Resalable Returns: A Single Period Model And Case Study”, *European Journal of Operations Research*, 2006, 169, 81-96.

[11] Pochampally, K. K., Nukala, S. and Gupta, S. M., “*Strategic Planning Models for Reverse and Closed-loop Supply Chains*”, CRC Press, Boca Raton, Florida, ISBN: 9781420054781, 2009.

[12] Shih, W., “A Note On Bayesian Approach To Newsboy Inventory Problem”, *Decision Sciences*, 2001, 4, 184-194.

[13] Thierry, M., Salomon, M., van Nunen, J., and van Wassenhove, L., “Strategic Issues in Product Recovery Management”, *California Management Review* 37(2), 114-135, 1995.

[14] Vlachos, D. and Dekker, R., “Return Handling Options And Order Quantities For Single Period Products”, *European Journal of Operations Research*, 2003, 151 (1), 38-52.

[15] Winkler, R. L., *An Introduction to Bayesian Inference and Decision*, Probabilistic Pub., Gainesville, FL., ISBN: 0964793849, 2003.

ANALYSIS OF REMANUFACTURING POLICY WITH CONSIDERATION FOR RETURNED PRODUCTS QUALITY

Kenichi Nakashima,Osaka Institute of Technology, Osaka, Japan, +81(6)-6954-4788, nakasima@dim.oit.ac.jp
Surendra M. Gupta, Northeastern University, Boston, MA 02115, (617)-373-4846, gupta@neu.edu

ABSTRACT

This paper deals with the product acquisition control problem and considers returned product quality with two types of classes. The system includes the flow of the product returns from customers to the factory as well as the forward flow of the sales. We formulate the acquisition problem together with product quality and stochastic demand using the Markov decision process. A numerical example is given to show the implementation of the methodology.

INTRODUCTION

From a macro-level perspective, value propositions for an organization or the industry in which the organization operates, including reverse supply chain strategies (reuse, repair, refurbish, remanufacture, retrieve parts or cannibalize components, recycle, scrap, and redesign returned products) and effective operations of reverse supply chains (handle returns, sort returns by value and ease of remanufacture), may augment organizational competitiveness. It is widely believed that the continuous growth in consumer waste in recent years has seriously threatened the environment. According to the US Environmental Protection Agency (EPA), in 1990, the amount of waste generated in the USA reached a whopping 196 million tons, up from 88 million tons in 1960s [13]. Wann [16] reported that an average American consumes 20 tons of materials every year. To ease such burden on the environment, many countries are contemplating regulations that force manufacturers to take back used products from consumers so that the components and materials retrieved from the products may be reused and/or recycled. For example, Germany has passed a regulation that requires companies to remanufacture products until the product is obsolete. Japan has passed similar legislation requiring design and assembly methodologies that facilitate recycling of durable goods [5]. Comparable regulations are also being considered in the United States. The two legislative acts that are expected to pass within the next few years in the U.S. are the Automotive Waste Management Act (which will enforce the complete reclamation of automobiles) and Polymers and Plastic Control Act (which will enforce the complete reclamation of polymers and plastics) [4].

In our research, we take a closed-loop supply chain perspective, rather than just a forward supply chain perspective, of a company's value chain. In such a closed-loop system, the reverse supply chain portion can create synergies with elements of the forward supply chain to form an integrated value chain environment. In this paper, we examine the optimal product ordering policy with consideration for the returned product quality. The system includes the flow of the product returns from customers to the factory as well as the forward flow of sales. We formulate this control problem together with product quality and stochastic demand using the Markov Decision Process [15]. We consider a single remanufacturing production process that produces a single item product. The finished products are stocked in the factory and are used to fulfill customer demand from outside. The product is produced using a returned product that belongs to either class 1 or class 2 quality. Each class has different acquisition cost, different remanufacturing cost and different delivery lead time. Therefore, the decision maker has to control two kinds of inventories for the retuned products.

The system is composed of the state that denotes the inventory levels of two quality classes of the returned products, the transition probabilities between states under a policy and the costs associated with the transitions. In this model, we control the numbers of each type of returned products: one is of high quality (class 1) while the other is of lower quality (class 2). We also consider the priorities for the use of the two types of products. Using Markov decision model [12], we can obtain the optimal ordering policy that minimizes the expected average cost per period. A numerical example is considered to illustrate the property of the control policy.

LITERATURE REVIEW

We present a brief review of the literature in the area of product recovery modeling of remanufacturing systems with stochastic variability.

Brennan et al. [1], Gungor and Gupta [5], Ilgin and Gupta [6] and Moyer and Gupta [9] reviewed the literature in the area of environmentally conscious manufacturing and product recovery. Minner [8] pointed out that there are the two well-known streams in product recovery research area. One is stochastic inventory control (SIC) and the other is material requirements planning (MRP). In this paper, we confine ourselves to SIC.

Cohen et al. [2] developed the product recovery model in which the collected products are directly used. Inderfurth [7] discussed effect of non-zero leadtimes on product recovery. Muckstadt and Isaac [10] dealt with a model for a remanufacturing system with non-zero leadtimes and a control policy with the traditional (Q, r) rule. Van der Laan and Salomon [14] suggested push and pull strategies for the remanufacturing system. Guide and Gupta [3] presented a queueing model to study a remanufacturing system. Nakashima et al.[11] considered a product recovery system with a single class product life cycle. In earlier research, existing models of reverse supply chains in the literature assume a constant quality of returned product. Moreover, product returns were assumed to have constant lead-times and constant cost for remanufacturing, which are not meaningful in industry settings. Addressing stochastic variability of quality of product returns can allow companies to realize additional profitability.

MODEL DESCRIPTION

We formulate a product aquisition system with stochastic variability using a discrete time Markov decision model. We consider a single process that produces a single item product. The finished products are stocked in the factory and are used to fulfill customer demand. The product is produced using a returned product that belongs to either class 1 or class 2 quality. Each class has different acquisition cost, different remanufacturing cost and different delivery lead time. Therefore, the decision maker has to control two kinds of inventories for the retuned products.

Figure 1 shows the product aquisition system in a remanufacturing environment. Remanufacturing preserves the product's identity and performs the required disassembly and refurbishing operations to bring the product to a desired level of quality at some remanufacturing cost. The number of products produced using normal manufacturing in period *t* is denoted by *P(t)*. All production begins at the start of a period and all products are completed by the end of the period. Product demand is independent and identically distributed (i.i.d) with mean *D*. The process produces the products using the recovered products that are supplied by two different suppliers with each own acquision cost. It is assumed that the leadtime of the part delivery is one. We use the following notations.

$I_n(t)$:inventory of class *n* (*n*=1,2) at the beginning of period *t*
$O_n(t)$:ordering quantity of class *n* at the beginning of period *t*
k_n :action as ordering part of class n （$k_n = O_n(t)$）
$D(t)$:demand in period *t*
a_n :acquisition cost per unit part for supplier *n*
h_n :holding cost per unit part supplied by supplier *n*
c_n :remanufacturing cost using part class *n*
$P_n(t)$:production quantity using pert class *n* in period *t*
C_b :backlog cost
L_1 and L_2 : the respective lead times for two classes of used products
p : the selling price of final products recovered (same for either quality class),

The state of the system is denoted by

$$s(t) = (I_1(t), I_2(t))$$

And, each inventory has maximum number $I_{1\max}$ and $I_{2\max}$. For the number of remanufacturing product: $P_n(t)$, the two resulting remanufacturing policies are provide as follows.

Class 1 Priority Policy:

$$P_1(t) = \min\{D(t), I_1(t)\},\ P_2(t) = \max\{0, D(t) - I_1(t)\}$$

Class 2 Priority Policy:

$$P_2(t) = \min\{D(t), I_2(t)\},\ P_1(t) = \max\{0, D(t) - I_2(t)\}$$

If Class 1 parts used prior to class 2

$$P_1(t) = \min\{D(t),[I_1(t)]^+\},\ P_2(t) = \max\{0,D(t) - [I_1(t)]^+,I_2(t)\},$$

where $[x]^+ = \max(0,x)$

If Class 2 parts used prior to class 1

$$P_1(t) = \max\{0,D(t) - [I_2(t)]^+,I_1(t)\},\ P_2(t) = \min\{D(t),[I_2(t)]^+\}$$

In regards to the action space, the numbers of orders for supplies are:

$$I_1(t+1) = I_1(t) + O_1(t-1) - P_1(t),\ I_2(t+1) = I_2(t) + O_2(t-1) - P_2(t)$$

Action spaces are shown by

$$K_1(s(t)) = \{0,\cdots, I_{\max 1} - I_1 - O_1(t-1)\} \text{ and } K_2(s(t)) = \{0,\cdots, I_{\max 2} - I_2 - O_2(t-1)\}.$$

Each action means that

$$k_1 = O_1(t), k_2 = O_2(t).$$

Transition Probability is

$$P_{s(n),s(n+1)}(k_1,k_2) = \begin{cases} \Pr\{D(t)=d\}, \; s(t+1)=\{I_1(t)+k_1-P_1(t), I_2(t)+k_2-P_2(t)\} \\ \text{Otherwise} \qquad\qquad ,0. \end{cases}$$

The expected reward is given by

$$r_{s(t)}(k) = pD(t) - \sum_{n=1}^{2}(k_n a_n + c_n P_n(t) + h_n I_n(t))$$

It is formulated as an average Markov decision process of time to maximum average profit, g:

$$g + v_i = \max_{k_1 \in K_1(i), k_2 \in K_2(i)} \left\{ r_i(k_1,k_2) + \sum_{j \in S} p_{ij}(k_1,k_2) v_i \right\}$$

We can calculate the stationary distribution of the system by solving a set of linear equations of the steady state distribution. We can then obtain the total expected cost per period using the above equation.

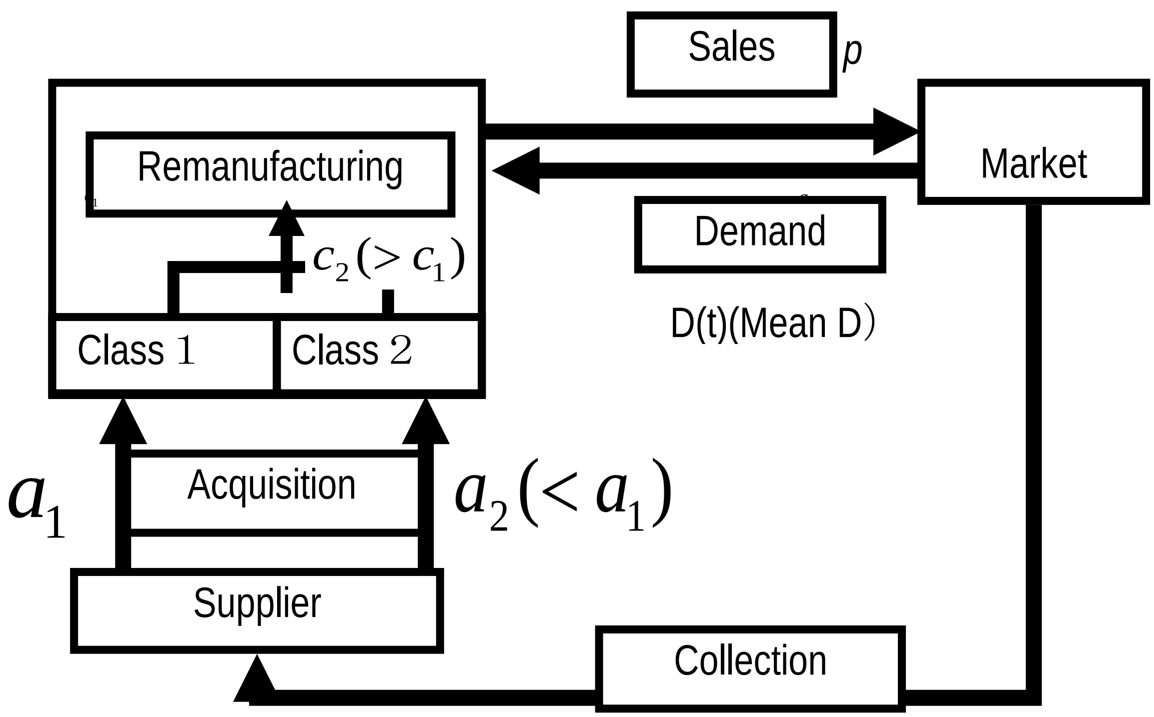


Fig.1: Remanufacturing Model

NUMERICAL RESULTS

In this section, we obtain the optimal ordering policy for a product acquisition system under stochastic demand.
The distribution of the demand is given by

$$\Pr\{D_n = D - \frac{1}{2}Q + j\} = \binom{Q}{j}\left(\frac{1}{2}\right)^Q, (0 \le j \le Q)$$

where D=3 and Q is an even number and the variance(σ^2) is $Q/4$. We can obtain the expected average reward per period under the steady state of the system.

$I_{1\max} = 5$; $I_{1\min} = 0$; $I_{2max} = 5$; $I_{2\min} = 0$; average D = 2.0, and decentralization = 1.0.

The cost parameters are: $(c_1,c_2) = (2, 4)$; $(a_1,a_2) = (2, 1)$; $(h_1,h_2) = (1, 1)$. Also, k_2 is assumed as fixed ordering system and considered as I_{max}, average profit computes to be 17.097. The optimal purchase policy for (c_1= 2, c_2 = 4) is provided in Table 1.

Table 1: Optimal Control Policy

Average demand 2	decentralization 1.0	g = 17.097	
(I_1, I_2)	k_1	(I_1, I_2)	k_1
(0,0)	5	(3,0)	2
(0,1)	5	(3,1)	2
(0,2)	4	(3,2)	2
(0,3)	1	(3,3)	2
(0,4)	1	(3,4)	0
(0,5)	1	(3,5)	2
(1,0)	4	(4,0)	0
(1,1)	4	(4,1)	0
(1,2)	1	(4,2)	0
(1,3)	1	(4,3)	0
(1,4)	1	(4,4)	0
(1,5)	1	(4,5)	0
(2,0)	3	(5,0)	0
(2,1)	0	(5,1)	0
(2,2)	3	(5,2)	0
(2,3)	3	(5,3)	0
(2,4)	3	(5,4)	0
(2,5)	3	(5,5)	0

CONCLUSION

We formulated the acquisition problem together with two types of product qualities and stochastic demand using the Markov Decision Process in a remanufacturing system. Numerical results illustrated the optimal ordering policy that maximized the expected average profit per period for the product acquisition system with different kinds of quality classes.

REFERENCES

[1] Brennan, L., Gupta, S. M. and Taleb, K. N., 1994, Operations Planning Issues in an Assembly/ Disassembly Environment, *International Journal of Operations and Production Management*, 14(9), 57-67.

[2] Cohen M.A., Nahmias S., Pierskalla W. P., 1980, A dynamic inventory system with recycling. *Naval Research Logistics Quarterly*, 27, 289-296.

[3] Guide Jr., V. D. R. and Gupta, S. M., 1999 (Mar. 1-3), A Queueing Network Model for Remanufacturing Production Systems, *Proceedings of the Second International Seminar on Reuse*, Eindhoven, The Netherlands, 115-128.

[4] Guide Jr., V. D. R., Srivastava, R. M. and Spencer S., 1996, Are production systems ready for the green revolution. *Production and Inventory Management Journal*, 37, 70-76.

[5] Gungor, A. and Gupta, S. M., 1999, Issues in environmentally conscious manufacturing and product recovery: a survey. *Computers and Industrial Engineering*, 36, 811-853.

[6] Ilgin, M. A. and Gupta, S. M., 2010, Environmentally Conscious Manufacturing and Product Recovery (ECMPRO): A Review of the State of the Art. *Journal of Environmental Management*, 91, 563-591.

[7] Inderfurth K., 1997, Simple optimal replenishment and disposal policies for a product recovery system with lead-times. *OR Spektrum*, 19, 111-122.

[8] Minner S., 2001, Strategic safety stocks in reverse logistics supply chains, *International Journal of Production economics*, 71, 417-428.

[9] Moyer, L. and Gupta, S. M., 1997, Environmental Concerns and Recycling/ Disassembly Efforts in the Electronics Industry, *Journal of Electronics Manufacturing*, 7, 1-22.

[10] Muckstadt J. A. and Isaac M. H., 1981, An analysis of single item inventory systems with returns. *Naval Research Logistics Quarterly*, 28, 237-254.

[11] Nakashima, K., Arimitsu, H., Nose, T. and Kuriyama, S., 2002, Analysis of a product recovery system: *International Journal of Production Research*, 40, 3849-3856.

[12] Nakashima, K., Arimitsu, H., Nose, T. and Kuriyama, S., 2004, Optimal control of a remanufacturing system: *International Journal of Production Research*, 42, 3619-3626.

[13] Nasr, N., 1997, Environmentally conscious manufacturing. *Careers and Engineer*, 26-27.

[14] van der Laan, E. A. and Salomon M., 1997, Production planning and inventory control with remanufacturing and disposal. *European Journal of Operational Research*, 102, 264-278.

[15] Puterman, M. L. 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York (John Wiley & Sons).

[16] Wann, D., 1996, *Deep Design: Pathways to a Livable Future*. Washington: Island Press.

ADDRESSING THE DISPARITY IN DECISION THEORY THE GESTALT IMPERATIVE

Samuel Kovacic, Old Dominion University, (757) 412-7113, skovacic@odu.edu,
Andres Sousa-Poza, Old Dominion University, (757) 412-7135, asousapo@odu.edu,

ABSTRACT

Decision science has provided methods for making decisions in a myriad of fields, but none as different as a multidisciplinary field. Multidisciplinary fields assume sufficient overlap exists between the disparate disciplines for effective decision making; however the gaps generally become more obscured, increasing uncertainty and complicating the decision process. This paper explores the nature of multidisciplinary problems, and the challenges it poses on current decision processes within complex situations. Additionally this paper suggests that complex situations are antithetical to mainstream interpretations for decision theory and concludes with an axiom in the form of an imperative for these traditional interpretations.

Keywords: complexity, gestalt, decision, multidiscipline

INTRODUCTION

Multidisciplinary Problems

Decision making in multi-disciplines potentially suffer from Type III error, solving the wrong problem precisely [Mitroff, 1998]. Multi-disciplinary study is an interpretation generated by combining differing disciplines aligned towards a common problem with the intent of generating a single perspective for greater understanding. The uncertainty generated by this merger, whether from the integration of two disciplines or the paradigmatic tension generated from multiple perspectives, seems to hang like a shroud over multidisciplinary fields of study. Uncertainty lies within these gaps and with it emergent and dynamic properties that over time constantly change the nature for how the problem is framed. Combining disparate disciplines, implying that a meta-type discipline would emerge, may have caused more uncertainty than certainty and generates debate on whether it contributes towards generalizable knowledge.

It is not contended that multidisciplines deal with high levels of uncertainty and its implications such as emergence, non-ergodic, and non-monotonic conditions that are open to multiple interpretations. Rather, it is contended that the decision making methods for dealing with these conditions have derived few benefits by the integration of the disciplines. Invariably, techniques for solving multidisciplinary problems generally are entrenched in one discipline and others are made to "fit" for a satisficing solution. Each discipline is immersed within its own lexicon and axioms, which are either subsumed by the other discipline or obviated entirely

Mainstream decision processes rely on ergodic and monotone conditions for the most effective decision choices. Although there is a movement for non-probabilistic approaches the countering argument is that these new approaches provide no better results than the probabilistic approach – reinforcing the penchant for a type three error – the approaches are not sufficient for solving problems found in complex situations.

The nature of complex situations

It is uncertain how many uses the word 'complexity' has found in the English language; as such it would be remiss not to frame how complexity is being used in this paper.

> "A key difference between current cybernetics and complexity theory is the use of different epistemology. Complexity theorists use a realist epistemology and assume that complexity exists in an observed system, or perhaps in a computer model. Cyberneticians use a constructivist epistemology and assume that the system of interest is defined by the observer." [Umpleby/WMSCI webpage],

This theme for how complexity is perceived and dealt with resonates within the research communities of many research and academic institutes, Umpleby's comments are foundational in BarYam's New England Complex Systems Institute (NECSI) where research explores both facets unilaterally. Complexity, a major division within the Santa-Fe institute has taken a multidisciplinary collaboration approach, and the University of Michigan's Center for the Study of Complex Systems encourage research in nonlinear, dynamical, and adaptive systems. Additionally, complexity is an integral thread in many centers: the National Centers of System of Systems Engineering (NCSOSE) invokes complex systems in their mission statement as does the System of Systems Center of Excellence (SOSECE). Sousa-Poza [Sousa-Poza et al, 2006], however, opens a unique door and suggests that complexity is tied not only to the observer and how he perceives but that the observer, as a participant, is a major contributor to the complexity, insinuating both a pragmatic and fallible component to complexity. Sousa-Poza extends on this theme by describing complexity as …"the probability for an erroneous knowledge claim". There is no doubt that complexity plays a role in the decision process, but it is the complex situation described by Sousa-Poza that is adopted in this paper and provides the impetus for the gestalt imperative.

BACKGROUND

The interpretations of Decision Theory

Currently there are three interpretations for decision theory, the third, a recent addition, representing a maturity in the former interpretations garnered over time. Until recently conventional wisdom had decision theory interpreted two ways: normative (also referred to as prescriptive) and descriptive. These interpretations provide theory for not only how a decision should be made (normative) but also how a decision is actually made (descriptive). Normative interpretations are empirically based and lend themselves to quantified techniques. Specifics of this method include detailed experiments for data gathering and bounding to assure accuracy of the decision. Descriptive interpretations (while no less quantifiable) lends itself to qualitative methods, where interviews and surveys are typical data gathering techniques and relying on rational bounding methods for affecting a decision. Both methods place the decision maker outside the decision space for an effective decision, arguably allowing for an unbiased perspective of the problem. The third interpretation, however, places the observer inside the decision space relying on expertise as the dominant method for data gathering and a bounding construct. This method supports a near-real time component to the decision process, where bias and change is obviated through expertise, and by default highlighting the influence of time on a decision. To understand how this may be relevant to the decision process it is necessary to review how knowledge is generated for decision theory.

The most notable and first philosopher for decision theory was Condorcet who hypothesized three stages in the decision process: discussion, clarification, and choice, [Condorcet, 1793, 1847]. Condorcet posited the intransitive nature of decision making, a very enlightened position, suggesting decisions were temporally bounded. Dewey veered off from Condorcet's theory and postulated five consecutive stages: a felt difficulty, the definition of the character of that difficulty, suggestion of possible solutions, evaluation of the suggestion, and further observation and experiment, [John Dewey's, 1910], this segue from Condorcet's, potentially process primed approach, along with Taylor's science management [Frederick W Taylor, 1911] integrated well with the dominant influence that sciences was having on the discipline, the demand for quantifiable knowledge as a basis for decisions and dominated research in the decision sciences to-date. Simon modified Dewey's work to facilitate decision within organizations and proposed three phases: intelligence, design and choice, [Herbert Simon, 1960]. Brim further added to Dewey's work to: Identification of the problem, obtain information produce possible solutions, evaluate solutions, and select strategy, [Brim et al., 1962]. Minzberg introduced a radically different perspective to the decision process migrating from the more sequential process to a simultaneous process. Minzberg suggested that the decision process was a parallel process rather than a sequential process and parsed his phases aligned with Simon's work: identification, development, and selection, [Mintzberg, Raisinghani, and Théorêt, 1976]. The distinction for Minzberg being that each phase is a routine and that their progression through the process is circular rather than linear.

The next paradigmatic change occurred in the late 1980s with a more contextually influenced approach. Naturalistic Decision Making (NDM) advanced the decision approach with their self proclaimed paradigmatic shift: making the observer the center of the decision space [Klein, Orasanu, Calderwood, & Zsambok, 1993]; [Lipshitz, Klein, Orasanu, and Salas, 2001]; [Hoffman,1995]; [Lipshitz, [2001]. Time pressure, uncertainty, ill-defined goals and high personal stakes were the contextual factors influencing decisions in crisis [Klein et al, 1993] and making the expert key for effective decision making. The idea that context must be established from within the domain rather than an empirically or rational extrapolation was a radical departure for the decision sciences, however, the method was exclusively for time sensitive decisions.

The common denominator for each of these interpretations is a phenomenological based construct absent Condorcet's original contribution, the dimension of time. Klein et al had made a major step towards the role of including the observer into the domain; however, it appeared the paradigm had not shifted sufficiently to break from the stronghold of a phenomenological interpretation.

By the time the third interpretation was introduced it became apparent that a foundational component was missing in the equation: the penchant for normative or prescriptive or naturalistic interpretations of decision theory in a substantive approach is absolute. However with the introduction of the third interpretation the idea that the stages of a decision process were not just sequential or parallel but based in a situation is important. The decision making stages went from a stochastic construct to a situational one defined by: situation type, suitable goals, courses of action, and relevance. This was the last piece of the puzzle to evolve the decision process under a new paradigm one that suggest that something lies beyond the known and known unknowns but into the realm of the unknown unknowns influenced most by the participant. The paradigmatic imposition of time insists that a decision must accommodate for change, uncertainty, and dynamic emergence to truly be effective.

DISCUSSION

The Basis for the Gestalt Imperative

In the evolution of decision sciences, three overarching disciplines have evolved. The evolution of each discipline was predicated by omissions in prior disciplines. Each discipline acts as a complement to the others, but is also complementary in nature. Bohr advocated that:

> *"evidence obtained under different experimental conditions cannot be comprehended within a single picture, but must be regarded as complementary in the sense that only the totality of the phenomena exhausts the possible information about the objects."* [Neils Bohr, 1949]

The mind-independent assertions of the positivist, and the mind-dependent assertions of the constructivist, are bounded by the paradigmatic imposition of the worldview or discipline of an observer. Holistic approaches are fundamentally contained within a discipline. The ability to function in a multidisciplinary manner is contingent on the degree to which shared awareness or understanding can be established from perspectives that are derived from different disciplines. In simple conditions where the axiomatic limitations of complementary perspectives are not challenged, a high degree of shared awareness is possible. The necessity of a multidisciplinary approach is, however, obviated by the correspondence of the conclusions that may be drawn by any discipline.

In complex conditions, complementary perspectives take on a polar or dual nature, in which the establishment of a common or shared awareness becomes increasing difficult if not impossible. It is, however, in such situations, where for example under conditions of emergence a Particularistic perspective finds no correspondence with a Universalistic perspective, that multidisciplinary approaches are *relevant*.

The necessity of simultaneously maintaining multiple decision disciplines can only be argued from the position of their indispensability to address a [complex] problem. For this, we must first establish the limitation of an orthodox (single discipline) position.

Within a discipline, the bounding of a problem is dictated by the principles and axioms that underlie the discipline itself. This act of bounding, however, not only influences the perspective(s) that can be supported, but the very manner that the problem and reality are perceived. The discipline in this sense will become "the hammer that makes everything look like a nail". The highest degree of comprehensibility will be marked by the nature of the bounding. Optimizing within this bound will maximize the understanding that is generated by a perspective, but will, based on the theory of complementarity [Rosenfeld, 1961], become increasingly polarized and inaccessible by other perspectives. In the problems where a satisfactory solution is identified within the comprehensibility that a perspective can provide, an orthodox position is warranted. If such a solution is however not possible within the constraints imposed by the comprehensibility of a perspective, adopting an alternate perspective, as is the case in reframing, might be possible. For truly complex problems any perspective will provide a local perspective, but will be unable to generate a sufficiently global construct to generate a suitable basis for further action. Thus, a paradoxical condition is set where multidisciplinary approaches can be enacted where they are not required (simple problems), and become impossible to adopt where they are necessary (complex problems).

A foundational concept of multidiscipline is *gestalt*. It implies that wholeness can never be recognized within the disparate disciplines that attempt to study the whole – they are too embedded within their own traditions. Instead emphasis is placed on the concept of '*meta'* [OED, 2008], or universal as suggested by Plato; more comprehensive concepts that convey the idea of wholeness to its parts; either concept holds true for decision making. The gestalt imperative is consequently proposed as a canonical set of principles for decision making; its purpose is to shape a decision space as a multidisciplinary whole for action.

Inclusivity Principle: to accommodate complementary perspectives within the decision process.

Non-Orthodoxy Principle: The Gestalt Imperative must have the mechanism embedded within it to provide the ability to let the gestalt influence the perspective without undue onus from kickback of the polar orthodoxy on the gestalt of the problem.

Pragmatic Principle: A practical confluence where the singular or monotone decision effect becomes nearly transient is mitigated by the participatory nature of the decision maker.

These principles provide the conduit to affect a gestalt imperative by providing the necessary feedback into the whole via an algedonic loop. The tension between the bi-polar perspectives is mitigated by the pragmatic nature of the naturalist feedback loop as well as the response from the gestalt (fig 1). The significance of the gestalt imperative is that it allows for viewing a problem not only phenomenologically (through observation) but noumenoloigcally (through participation) accepting the nuance that uncertainty, which monotone perspective attempt to remove, is better accepted than extracted or obviated.

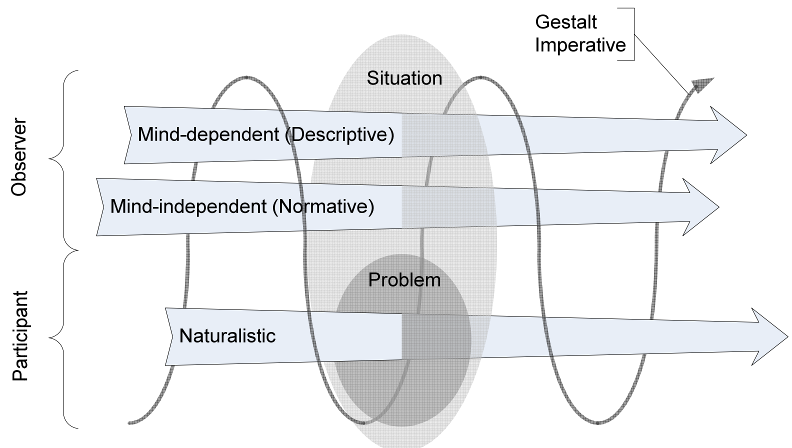


Figure 1: The Gestalt Imperative

The Influence of the Gestalt Imperative on the Stages of Decision Making

Sun Tseng (para-phrased) once stated: “the victorious army is the one that sets the condition for victory and then pursues the condition(s), while the loosing army defines and pursues victory”. Loosely interpreted it is more important to understand where you are going than how you get there. This is the paradigm that is adopted for integrating all three interpretations and adopting a more pragmatic approach for decision making. What Sun Tseng is referring to is a process philosophy. [Nicholas Rescher, 2000] states: “Several, if not all, of the major elements of the ontological repertoire (God, Nature as a whole, persons, material substances) are best understood in process terms…

> …A process philosopher, then, is someone for whom temporality, activity, and change-of alteration, striving, passage, and novelty-emergence-are the cardinal factors for our understanding of the real.”

Nicolas Rescher aptly highlights the need for a process approach as well as the traditional substantive approach for decision making, especially one where the decision is relegated to a person. This would suggest that a reinterpretation of the approach highlighting the conditions for a decision rather than the approach is required. Under this modality an assessment of the stages of a decision might be re-interpreted as: *a) truth, the acceptance that reality is greater than the purview of any one perspective, b) plentitude, concedes that there is always one best possibility for any given situation at any given time, and, c) action, sets the metrics for the process as well as the algodonic loop back for truth,* each stage having a reciprocal relationship with the other. The reciprocal nature of the process encourages understanding at the same time fosters an intuitive/tacit flavor critical for maintaining the necessary temporal dimension of the individual while maintaining a coherent relationship with atemporal aspect of the whole, key for the interpretation of decision theory.

CONCLUSION

Is there merit in the Gestalt Imperative? Difficult decisions are made daily, yet there is a perceived limitation in decision making, sufficiently that the Department of Homeland Security, Systems Engineering Solutions Inc (SESI), and Old Dominion University (ODU) have formed and funded a partnership to explore potential methods for making decisions in complex situations. A proposed method, Reverse Decision Making, will be vetted over the three year period that the partnership has been formed. It is anticipated that the Gestalt Imperative and more practically RDM will provide a means for decision making in complex situations. In summary: complexity, measured by the probability of an erroneous knowledge claim, imparts assumptions that not everything within any situation can be completely known and that a decision space must account for the known, known unknowns, and the unknown unknowns. Without a holistic interpretation accounting for the empirical, rational, and practical aspects of knowledge current interpretations of decision theory is fundamentally incomplete.

References available on request from Samuel Kovacic.

SURVIVAL AND PROPORTIONATE GROWTH IN INNOVATION: FIRMS AND CITIES

Thomas F. Brantle, Stevens Institute of Technology, (732) 872-7742,
thomas.brantle@stevens.edu

ABSTRACT

The survival and growth of firms and cities globally is investigated using the number and size of a firm's and city's constituent 'innovation' components. Innovation as characterized by patents, citations and inventor activity is used as a measure of firm and city size. Patents, citations, and inventor activity have been considered excellent proxies for the innovation process, technological advancement and resultant economic growth. It is shown that the growth rate distribution of firms and cities globally is well represented by Gibrat's weak law of mean proportionate growth and thus the expected growth rate of a given firm or city is independent of its size at the beginning of the examined period across time and using several size measures of innovation production and activity. Additionally it is seen that the survival rate also increases with firm or city size across all innovation size measures and time periods considered. These results are confirmed empirically. A direct connection is established between the structure of organizational and geographic growth as represented by innovation production and activity.

INTRODUCTION

Earlier research has hypothesized and demonstrated that the agglomeration and residential mobility of the population between different geographic locations is tightly connected to associated economic activity. Thus, a full understanding of organizational and geographic activity involvers understanding population mobility along with associated economic driving forces (e.g., innovation and technology). The firm and city size distributions are relevant indicators of industrial and population concentrations, respectively, quantities of interest to both technology and financial managers as well as antitrust and public policy decision makers. [4] [6] [11] [14] [15] [16] [17] [19] [20] [21] [23]

One of the most important issues in the industrial organization and population dynamics literature is the growth rate of both firms and cities, which is typically analyzed from the perspective of the level of agreement or divergence from Gibrat's Law of Proportionate Growth. Gibrat's law implies that the expected growth rate of a given firm or city is independent of its size at the beginning of the examined period. Or alternatively, it may be stated as the probability of a given proportionate change in size during a specified period is the same for all firms and/or cities in a given industry or region as applicable, regardless of their size at the beginning of the period. [5] [7] [12] [18] [22]

Two basic versions of Gibrat's law are typically considered: a 'weak' version which simply supposes that the mean growth rate is independent of size over time; while a 'stronger' version which usually hypothesizes that the variance of growth rates is also independent of size over time. Other issues sometimes considered from both theoretical and empirical perspectives include: entry rates and the shape of the growth rate distribution. This paper will focus on the first issue, that is, an examination of the weak version of Gibrat's law and the level of agreement or disagreement with the independence of mean growth rate vs. size over time, as well as the related issue of the survival rate of firms and cities vs. size over time. [5] [7] [12] [18] [22]

Empirical investigations of Gibrat's law typically estimate equations of the following type: $\log(S_t) = \alpha + \beta\log(S_{t-1}) + \varepsilon$ (i.e., where the underlying function is a power law of the form: $S_t = \alpha S_{t-1}^{\beta}$), where a firm or city's size is represented by S_t at time t and α is a constant term, representing the industry or region-wide growth trend. For Gibrat's law, the focus is on the value of the coefficient β. If firm or city growth is independent of size, then β takes on the value unity (i.e., $\beta = 1$). If β is smaller than one, then smaller firms or cities grow faster than their larger counterparts (e.g., regression to the mean). On the other hand, if β is larger than one, then larger firms or cities grow relatively more rapidly and there is an inclination toward industry or population concentration and monopoly concerning firms. [5] [7] [12] [18] [22]

Previous studies have differed as to the applicability of Gibrat's law to large firms and cities vs. the inclusion of increasingly smaller sized firms and cities. Early studies focused on large firms or cities, partly due to the availability of such data, and revealed either general agreement with proportionate growth or a slight tendency for larger sized firms to have higher growth rates. However, the majority of recent studies of firms have utilized larger and more diverse datasets, and have found values of β slightly lower than unity, which implies that on average, smaller firms seem to grow faster than larger firms. Studies of cities have generally been limited to larger cities and the largest metropolitan areas, and have shown general agreement with stable proportionate growth. [2] [9] [10] [21]

Conversely, these studies of growth rate conditioned on firm size are often affected by sample selection or sample attrition issues, as smaller firms have a high probability of exiting the industry (failure). This has been verified by several studies that indicate that the probability of firm survival (failure) increases (decreases) with firm size, and similarly the proportionate growth rate of a firm conditional on survival decreases with firm size. [2] [9] [10] [21]

Firm size has been typically measured by total assets, sales revenue, market value or number of employees and city size determined by population census or income figures. The firm and city size growth rates are relevant indicators of industrial and population concentrations, respectively, quantities of interest to both technology and financial managers as well as antitrust and public policy decision makers. [5] [7] [12] [18] [22] [23]

The survival and growth of economic activity, as represented technological innovation, is investigated across geographic locations (e.g., cities) and organizations (e.g., firms). Specifically, the survival rate and mean proportionate growth rate of firms and cities globally, as represented by innovation production and activity, is considered in this research. Innovation has long been seen as driving productivity and economic growth. Patents have long been explored and recognized as a very useful and productive source of data for the assessment of innovation development and technological advancement.

INNOVATION AND PATENT DATA

Patents provide a wealth of information and a long time-series of data about inventions, inventors, collaboration, prior knowledge, and the assigned owners and locations (e.g., firms and cities). Patents and the inventions they represent have several advantages as a technology indicator. In particular, patents and patent citations have long been recognized, and more recently invention collaboration, as a very rich and fertile source of data for studying knowledge, innovation and technological change. As such, they provide a valuable tool for public and corporate technology analysis, planning and policy decisions. [3] [8] [13]

Patents and patent citations constitute a documentation path for knowledge creation, technology transfer and innovation flows. When one patent cites another patent the citation indicates knowledge and information flowing from the cited patent to the citing patent. Therefore, the knowledge and innovation information made publicly available by the cited patent has not only flowed to the citing patent but has influenced both the citing patent's inventor(s) and the invention it represents. Two sets of measures are generally considered: forward-based measures that are derived from the relationship between a given patent and subsequent technology developments that build upon it, which are considered descendents; and backward-based measures that are derived from the relationship between a given patent and the body of knowledge that preceded it, which are considered antecedents. The supposition is that forward-based measures would be informative of the subsequent impact of research outcomes, whereas backward-based measures would be informative of the nature of the research. Forward citations or citations received measure originality and are an indication that the technology area is new or less-developed and the innovation is more likely to be significant and of greater value. Backward citations or citations made measure generality and are an indication that the technology area is well-developed and the innovation is more likely to be derivative and of less value. [13] [24]

The United States Patent and Trademark Office (USPTO) utility patent, citation and inventor raw data files are used in this study. [1] First, all the data is appropriately cleaned, filtered and reformatted for data analysis purposes. The data comprises detailed information on approximately 3.1 million U.S. patents granted between January 1975 and December 2005. Each patent contains highly detailed information on the invention itself, the technological area to which it belongs, the inventors, including their geographic location, and a company or organizational identifier. Additionally it includes the list of all citations made by patents issued between 1975 and 2005; about 21.6 million citations that either cites are cited by other patents. In addition the patent inventor data file contains approximately 6.6 million unique individual patent-inventor pairings, consisting of about 3.1 million patents and roughly 1.9 million individual inventors, from 1975 through 2005.

For each patent the geographic location (country, state and city) is determined by the "first" or "primary" inventor on the patent. The company or organizational identifier of each patent is known and recorded for approximately 84% of the set of patents, with the remaining 16% generally referred to as the "independent inventors" which includes small businesses, self-employed inventors and other small organizations. All these patents and inventors represent companies, organizations and individuals across both private industry and government organizations, with about 55% of patents produced and owned by U.S. entities and the remaining 45% spread across the globe and numerous countries.

To account for the patents associated with the “independent inventors” an additional organizational identifier is constructed. For each of these “independent inventors” patents its “first” inventor is considered equivalent to a small business owner or organizational head and assigned its own unique organizational identifier. Thus all patents produced under this individual are considered owned by this small business or organization. Similarly, all inventors associated with these independent inventor patents are considered employees or associates of this small business or organization. Thus, two organizational identifiers are identified and/or constructed, the first is the original assignee heretofore denoted as “Companies” and the second which includes the known original assignees plus the independent inventors will be referred to as “Firms”. Hence Firms includes the Companies as well as the independent inventor small businesses, self-employed, as well as other small organizations.

In addition, two versions of Firms and Companies are considered, the first is unrestricted by geographic boundaries and these are referred to as Global Firms and Global Companies, respectively. The second restricts Firms and/or Companies to either their State boundaries for U.S. owned patents and or its Country boundaries for non-U.S. owned patents, and these are referred to as Regional Firms and Regional Companies. Thus, the Regional Firm and Company identifier represents the local or regional presence of an organization while the Global Firm and Company identifier represents its worldwide or global reach.

Cities are identified and recorded for each Patent based on the city of the “first” or “primary” inventor and is considered the geographic location of the invention. While, individual inventor’s Cities are identified and recorded based upon the location and city of residence of the inventor at the time the patent was granted. Thus individual inventors may be associated with different Cities over time and their patenting production. Similarly, inventors may be associated with different Global Firms and Companies or Regional Firms and Companies over time and their patenting production.

The innovation size measure for all firms, companies and cities is tracked over 5-year periods, keeping track of firm, entry, growth, and exit, considered as no growth for each time period. Also, if a firm, company or city shows no growth (constant size) over consecutive 5-year periods then the next time it shows growth it is considered a new entry.

SURVIVAL AND PROPORTIONATE GROWTH – RESULTS DISCUSSION

The growth of firms and cities is considered over 5-year time intervals from 1975 to 2005 for several size measures of innovation. Firm and city innovation activity is measured by patenting and inventor production and activity, specifically the number of forward citations (cited patents, an output measure of originality, productivity and value), backward citations (citing patents, an input measure of generality, derivation and cost), total number of patents (an output measure of productivity and value), and the number of inventor instances associated with patent production (an input measure of human capital). Two definitions of a firm’s boundaries are considered, one which represents its global innovation activity and another restricting it to its country or state boundaries. Importantly all sizes of firms, companies and cities are considered, including the full recognition of independent inventors (e.g., small businesses and individuals) which represents roughly one-sixth of patents and invention activity.

Probability of Survival in Innovation

Figures 1, 2 and 3 plot the probability of survival vs. innovation size on double-logarithmic graphs for Global Firms and Companies, Regional Firms and Companies and All Cities, respectively, over 5-year time periods and for all four measures of innovation size. It is visually apparent that smaller firms, companies and cities demonstrate a higher exit or failure (no growth) rate than their larger counterparts. All firms, companies and cities also demonstrate parallel result patters for all innovation size measures with a noticeable rise in survival as past innovation size increases. The slight shift of the curves to the right for each successive 5-year period is simply attributable to the fact that surviving firms are always growing or increasing in size over time and not a unique difference between time-period dynamics. It should be noted that positive growth in number of backward citations, patents and inventor instances is associated with active and ongoing patent production. While, positive growth in number of forward citations indicates that a firm, company or city is still receiving value or rewards from previously produced patents.

Proportionate Growth in Innovation

Figures 4, 5 and 6 plot the slope or scaling exponent of the mean proportionate growth rate β for Global Firms and Companies, Regional Firms and Companies, and All Cities, respectively, over 5-year time periods, for all four measures of innovation size. There are two growth cases considered for each, the plots on the left consider only surviving firms, companies and cities and thus estimates the slope of the mean proportionate growth rate using only positive growth values. While, the plot on the right, considers the case including both surviving and exiting (or no growth and constant size) firms, companies and cities in estimating the slope or scaling exponent of the mean proportionate growth rate β. Note all OLS linear regression model fits estimating the mean proportionate growth rate parameter β yield an $R^2 \sim 0.98\text{-}1.00$.

In general the slope or mean proportionate growth rate β is between 0.90 and 1.10 for all firm, company and city growth innovation size scenarios considered. In particular when only surviving firms are considered then the mean proportionate growth rate slightly favors smaller firms, companies and cities growing faster (i.e., $\beta < 1$). However, this is conditioned on survival, where previously it was shown that smaller firms, companies and cities exit or fail (no growth) at a much higher rate than their larger counterparts. Now, in the case where all growth (surviving) as well as no growth and constant size (exiting or failing) firms, companies and cites is considered then it is clearly evident that larger firms, companies and cities demonstrate a higher growth rate than smaller entities (i.e. $\beta > 1$), as one might expect.

In addition, it is observed that for both the positive growth as well as the growth/no growth cases, forward citations (which measures productivity and greater value) demonstrates the greatest stability between the two cases varying and differing only slightly. While backward citations (which measures generality, derivation and cost of information) shows the greatest variability between the two cases, indicating a lower cost of information for surviving entities vs. failing entities. Number of patents (also a measure of productivity) and number of inventor instances (a measure of human capital cost) show similar and parallel behavior for all firm, company and city scenarios considered, varying more than the forward citations mean proportionate growth rates but less than the backward citations growth rates for these two cases.

FIGURE 1

Global Firms and Companies – Probability of Survival

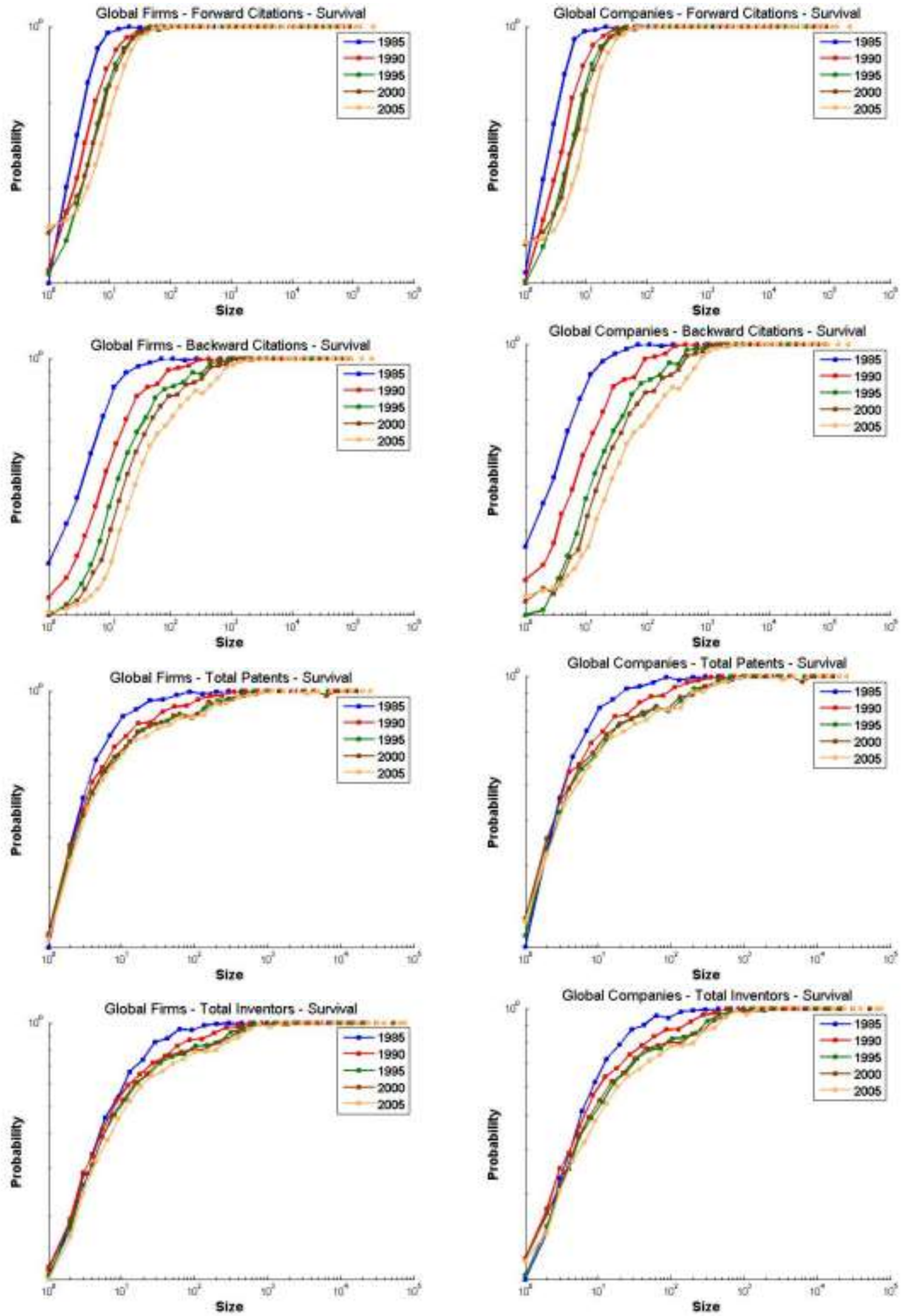


FIGURE 2

Regional Firms and Companies – Probability of Survival

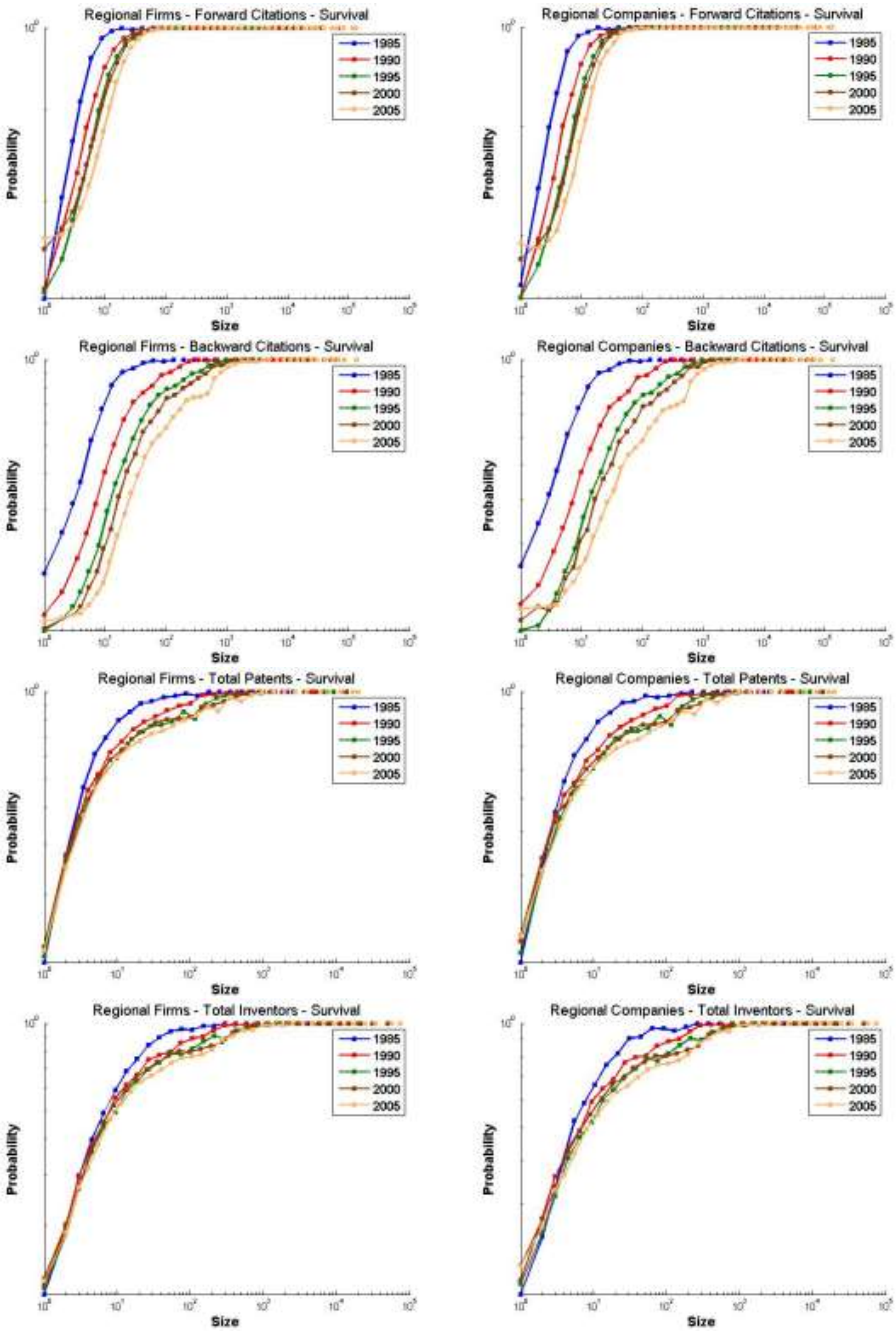


FIGURE 3

All Cities – Probability of Survival

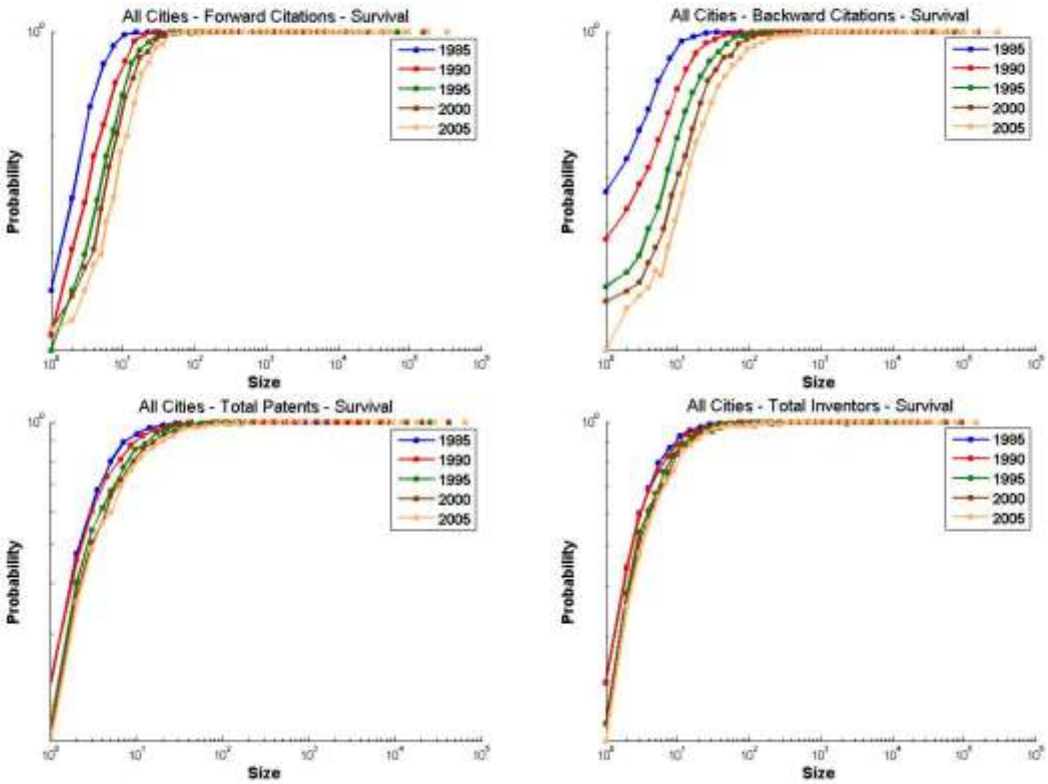


FIGURE 4

Global Firms and Companies – Proportionate Growth

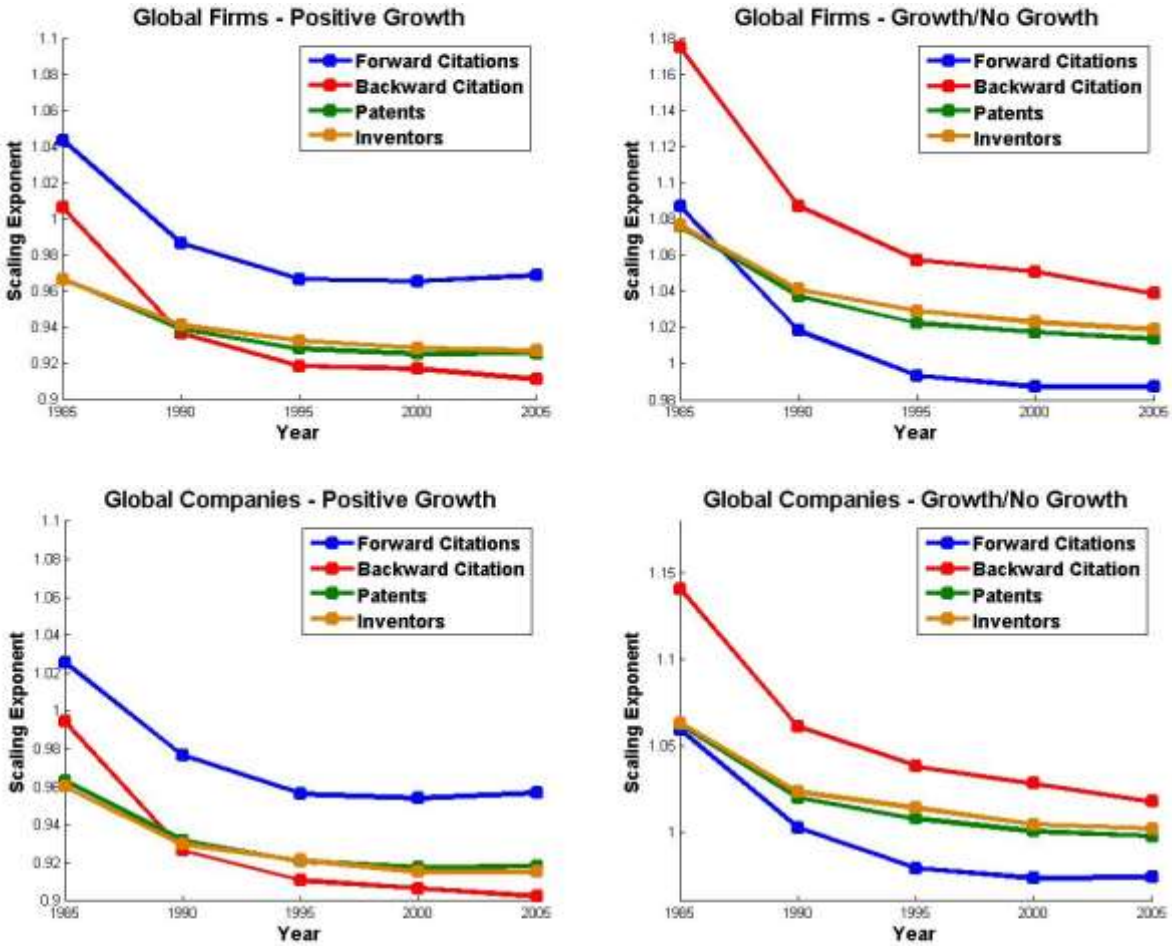


FIGURE 5

Regional Firms and Companies – Proportionate Growth

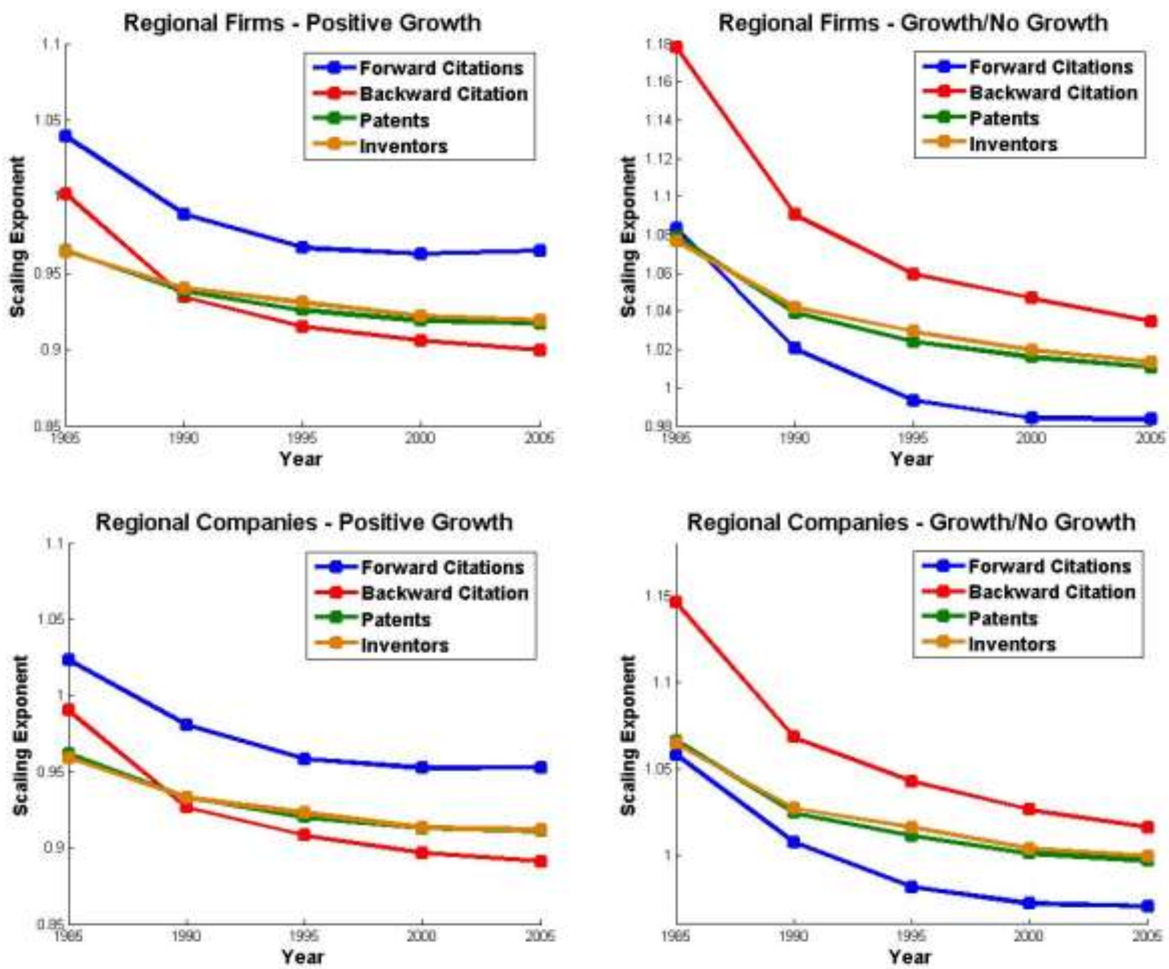
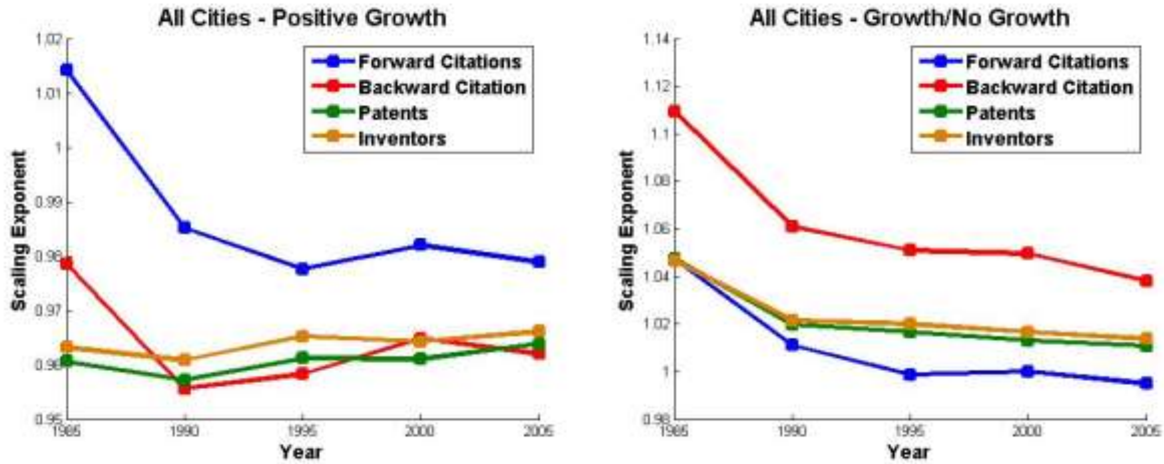


FIGURE 6

All Cities – Proportionate Growth



SUMMARY AND CONCLUSIONS

It is generally found that in all cases, across innovation size measures, the growth rates of firms and cities globally for innovation production and activity is well represented by Gibrat's weak law of mean proportionate growth, and hence the expected growth rate of a given firm or city is independent of its innovation size at the beginning of the examined period. Thus the probability of a given proportionate change in firm or city size during a specified period is the same for all firms and cities worldwide and across technology domains, regardless of their innovation size at the beginning of the period. Therefore, even though growth rates between different firms and cities vary substantially, there is no systematic pattern with respect to size, i.e., the underlying stochastic process is the same for all firms and cities. Consequently, with regards to innovation production and patenting activity the results indicate that larger firms, companies and cities on average do not grow faster or slower than smaller firms and cities.

Additionally it is seen that the survival rate also increases with firm and city size across all innovation measures considered. Thus, the probability of survival increases with firm and city size, and thus the proportional growth rate of a firm or city conditional on survival is decreasing in size. Although, this impact on the overall observed expected growth rate curve is minimal for firms and cities, i.e., mean proportional growth is still very much evident. Nonetheless, the results that firm or city survival increases with size is consistent with smaller firms or cities being more volatile (higher variability) and large firms and cities being more stable (lower variability) observed elsewhere.

What these results indicate is that there are two effects at work here in the size-growth relationship: larger firms and cities have generally equal growth rates as compared to their smaller counterparts; however they are also more likely to survive regarding their innovation production and patenting activity. Thus, evidence is provided of mean proportionate growth rate for firms and cities is independent of the definition of innovation size chosen. The existence of

Gibrat's law of mean proportionate growth is seen to be equally robust for firms and cities across several measures of innovation and size.

This research represents a unique and first time contribution to the economic, organization, innovation, and firm and city growth rate distribution literature as it represents firms and cities globally, includes not only the largest firms and cities but also the smallest firms and cities in terms of innovation production and activity, and importantly is the first time the survival rate and growth rate for firms and cities has been so described and analyzed for innovation production and activity, i.e., patenting and inventor activity and production. It also indicates that the survival and growth of innovation production and activity for firms and cities, i.e., across geography and organization, is robust and follows a parallel pattern and structure. Thus, a direct connection is established between the structure of organizational and geographic survival and growth as represented by innovation production and activity.

REFERENCES

[1] *United States Patent and Trademark Office (USPTO) Technology Assessment and Forecast (TAF) Database*. 2005, United States Patent and Trademark Office (USPTO), Electronic Information Products Division: Alexandria, VA 22313.

[2] Bottazzi, G. and A. Secchi, *Explaining the Distribution of Firm Growth Rates.* The Rand Journal of Economics, 235-256. **37**(2, Summer 2006): p. 235-256.

[3] Cohen, W.M. and S.A. Merrill, eds. *Patents in the Knowledge-Based Economy*. 2003, The National Academic Press: Washington, DC.

[4] Cordoba, J.-C., *On the Distribution of City Sizes.* Journal of Urban Economics, 2008. **63**: p. 177-197.

[5] Eeckhout, J., *Gibrat's Law for All Cities.* American Economic Review, 2004. **94**(5): p. 1429-1451.

[6] Fujita, M., P. Krugman, and A.J. Venables, *The Spatial Economy:Cities, Regions and International Trade*. 2001, Cambridge, MA: The MIT Press.

[7] Geroski, P.A., *The Growth of Firms in Theory and Practice*, in *New Directions in Economic Strategy Research*, N. Foss and V. Malinke, Editors. 2000, Oxford University Press: Oxford. p. 168-186.

[8] Griliches, Z., *Patent Statistics as Economic Indicators: A Survey.* Journal of Economic Literature, 1990. **28**(4): p. 1661-1707.

[9] Hall, B.H., *The Relationship Between Firm Size and Firm Growth in U.S. Manufacturing Sector.* Journal of Industrial Economics, 1987. **35**: p. 583-606.

[10] Hart, P.E. and S.J. Prais, *The Analysis of Business Concentration: A Statistical Approach.* Journal of the Royal Statistical Society, 1956. **119**: p. 150-191.

[11] Henderson, J.V., *Urbanization and Growth*, in *Handbook of Economic Growth*, P. Aghion and S.N. Durlauf, Editors. 2005, Elsevier B.V.: Amsterdam.

[12] Ijiri, Y. and H.A. Simon, *Skew Distributions and the Sizes of Business Firms*. 1977, New York: North Holland.

[13] Jaffe, A.B. and M. Trajtenberg, eds. *Patents, Citations, and Innovations: A Window on the Knowledge Economy*. 2002, MIT Press: Cambridge, MA.

[14] Krugman, P., *Development, Geography and Economic Theory*. 1995, Cambridge, MA: MIT Press.

[15] Krugman, P., *Confronting the Mystery of Urban Hierarchy.* Journal of the Japanese and International Economics, 1996. **10**: p. 399-418.
[16] Krugman, P., *The Self Organizing Economy*. 1996, Malden, MA: Blackwell Publishers Ltd.
[17] Lucas, R.E.J. and E. Rossi-Hansberg, *The Internal Structure of Cities.* Econometrica, 2002. **70**(4): p. 1445-1476.
[18] Mansfield, E., *Entry, Gibrat's Law, Innovation and the Growth of Firms.* Americvan Economic Review, 1962. **52**: p. 1023-1051.
[19] Porter, M.E., *Clusters and the New Economics of Competition.* Harvard Business Review, 1998. **78**: p. 77-90.
[20] Rosenthal, S.S. and W.C. Strange, *Geography, Industrial Organization, and Agglomeration.* The Review of Economics and Statistics, 2003. **85**(2): p. 377-393.
[21] Simon, H.A. and C.P. Bonini, *The Size Distribution of Business Firms.* American Economic Review, 1958. **48**(4): p. 607-617.
[22] Sutton, J., *Gibrat's Legacy.* Journal of Economic Literature, 1997. **XXXV**: p. 40-59.
[23] Sutton, J., *Technology and Market Structure - Theory and History*. 1998, Cambridge, MA: MIT Press.
[24] Trajtenberg, M., R. Henderson, and A.B. Jaffe, *University versus Corporate Patents: A Window on the Basicness of Innovation.* Economics of Innovation and New Technology, 1997. **5**: p. 19-50.

AGAINST ALL ODDS:
A STRATEGIC ANALYSIS OF THE FALL OF HONG KONG, 1941

Ed Chung, chunge@etown.edu
Cristina Ciocirlan, ciocirlanc@etown.edu
Bingye Mu, mub@etown.edu
(all at Elizabethtown College, Business Dept., Elizabethtown, PA 17022)

ABSTRACT

The objective of this paper is to employ a historical case event in an analysis of competitive strategies. We borrowed from a model from extant literature which utilized a similar historical analysis approach in a study of strategic decision-making. Using the (unsuccessful) defence of Hong Kong in World War II as the historical case, we seek first to apply Chung and McLarney's (1999) model in our analysis, and then extend the model so as to better handle the unique sequence of events that took place in 1941. Our study not only modifies and extends the model, but also contributes to the literature by augmenting the validity of previous case research.

Keywords: Case study, decision model, heuristics, biases

INTRODUCTION

The objective of this paper is to employ a historical case event in our analysis of competitive strategies. Specifically, we are interested in the decisions that were made, and how these impacted the outcomes. To do so, we borrowed from a model from extant literature which utilized a similar historical analysis approach in a study of strategic decision-making.

Chung and McLarney (1999) drew lessons from a historical analysis of the Battle of Midway (a watershed event that changed the course of the war in the Pacific during the Second World War), and developed an extension of Daft and Weick's (1984) conceptualization of organizations as "interpretation systems." According to Chung and McLarney's model, three key components interact in a strategic decision-making process: First, selective filtering takes place during data collection; second, the congeries of leadership style, organizational culture, and structural considerations shape how the collected data are interpreted; and third, the actual decision and its implement feed back into a learning process which then loops back to inform each of these three components. Using this model, Chung and McLarney demonstrated the strategic mistakes that were made and how these impacted the outcome of the battle. They urged fellow investigators to examine "other similarly turbulent decision-making environments" to see if comparable findings can be obtained. Indeed, they wrote "it may also be fruitful to test our model under different situations and see how we may better refine the model itself" (p.245).

This paper attempts to build on previous work by analyzing a similarly turbulent decision environment. Using the (unsuccessful) defence of Hong Kong in World War II as the historical case study, we seek first to apply Chung and McLarney's model in our analysis, and then extend the model so as to better handle the unique sequence of events that took place in 1941 in Hong

Kong. Our study not only modifies and extends the model, but also contributes to the literature by augmenting the validity of previous case research.

HISTORICAL BACKGROUND

> All the time we knew, or at least I did, that we were fighting a losing battle. It was so different to what I'd expected. We didn't stand a chance. (Kelly, 1942)

The above quote was taken from the diary of Inspector Fred Kelly of the Hong Kong Police Force, who was interned in the Stanley internment camp (in Hong Kong) for the duration of the Japanese occupation. His introspection confirms what many other writers have noted, that the defence of Hong Kong, Churchill's empty promises notwithstanding, was untenable. Indeed, Lai (1999) noted that Churchill himself had decided early in 1941 that the colony could not be defended, and that reinforcements would be wasted. And yet, against all odds, the defenders of Hong Kong stood their ground far longer than anyone had envisioned, and exacted a heavy toll on the invading Japanese forces.

On December 8, 1941, eight hours after the attack on Pearl Harbour, Japan launched its invasion of the British Crown Colony of Hong Kong. The invading forces numbered more than 50,000 battle-hardened troops who had been fighting in China for the past four years, while the defenders totalled roughly 11,000 to 14,000 (estimates vary) green troops, many who had never fired a weapon in anger. Seventeen days later, after suffering extraordinarily high casualties, running out of supplies, and with no reinforcements in sight, the Allied Forces (an amalgamation of British, Canadian, Indian, and locals) garrison surrendered to the superior Japanese forces, thus begun three years and eight months of Japanese rule in the Colony. Major General J.N. Kennedy argued that the 17 days defence of Hong Kong sent a signal to the Japanese regarding the resilience of the Allied Forces and would give them a standard by which to judge their resistance in future battles, such as Malaya and others (Lindsay 2005).

The defending Allied troops suffered casualties well in excess of 30 percent during the brief conflict. Because of war-time (and subsequent POW camp) conditions, precise figures are unavailable. But according to the despatch the British commanding general (Major General Maltby) filed with the Imperial War Office after the war, some 4,500 of the defending troops were either killed, wounded, or missing in action. Those who survived did not have a much happier fate, for awaiting them were the dreaded POW camps courtesy of the Imperial Japanese Army. Many of those thus incarcerated did not survive.

And what of the invaders? The numbers are even murkier, though estimates have been put between 2,700 killed and wounded, to 27,000 killed and wounded. Banham (1999), citing sources from the Hong Kong News published immediately after the fall of Hong Kong, suggested that an estimated casualty of 8,000 would be reasonable. Not quite a pyrrhic victory, to be sure, but certainly one that did not come cheap.

A DECISION MODEL

Building on the Chung and McLarney's (1999) work, we add elements well-developed in the organizational behaviour literature, such as leader's personality in relation to risk perception, problem framing, ethics, team dynamics, and emotions. The model developed (Figure 1) here will explain how these elements had an impact the decisions made in the Hong Kong battle.

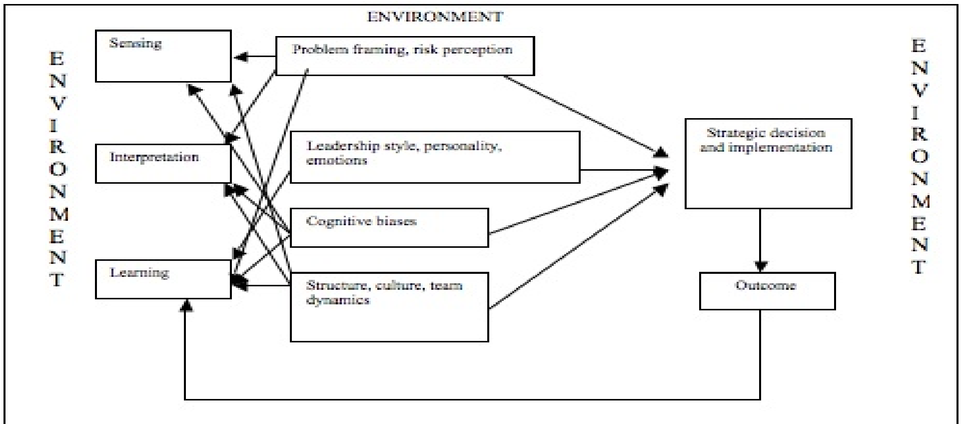


Figure 1. A decision-making model.

In our model, elements such as problem framing, leadership style, personality and emotions, cognitive biases and structure, culture and team dynamics affect all three phases of the decision-making process, sensing, interpretation and learning. The way in which the problem is formulated has an impact upon the way in which decision-makers make sense of their environment, but also on the way in which they interpret information and the way in which they learn. Similarly, cognitive biases are present in the data collection phase, as well as in the interpretation and learning phase. The process of decision-making is messier, more dynamic and claims no linearity between the three phases. By the time decision-makers make sense of their environment, the environment changes and needs to be assessed again, using, perhaps a different frame of reference, which may be "loaded" with different biases.

THREE DECISION POINTS

The rational decision-making model states that, faced with a certain problem, decision-makers should lay out all possible alternatives, develop criteria, rank these criteria according to importance, evaluate alternatives with respect to how well they meet these criteria and choose the optimum solution. It is a normative model which specifies how decisions should be formulated, rather than how they are actually formulated in practice (Knighton 2004). However, we will show how the rational decision-making model consistently breaks down in explaining decisions made at three major decision points during the battle.

The Decision to Defend Hong Kong

In 1938, the British officers and former governor considered the possibility of an "open city," meaning that the Japanese forces would be allowed to enter freely without a fight (Lindsay 2005). Approached rationally, the "open city" option would have been optimal. Yet this was not to be.

Our research indicates that the decision to fight was driven by considerations not necessarily rational, though important nonetheless. Such factors as political and morale considerations, leadership personalities, value and risk perceptions, faulty scanning, decision-making biases, and so on contributed to the fatal decision to defend the isolated outpost “at all costs.” Space does not allow us to detail all such problematics, but we offer here one example for illumination: Air Chief Marshal Sir Robert Brooke-Popham felt that the Japanese were vastly inferior to the British forces. He based his opinions on trivia such as the dirty uniforms of the Japanese, while disregarding the fact that the Royal Air Force had only five operational planes in theatre (Lindsay 2005). Logic such as this moved the defenders to their fateful rendezvous with history.

Decisions During the Battle

Several factors worked against the Allied Forces in their defence of the island. First, personal and organizational experience had an important impact upon the decisions made and, ultimately, on the outcome of the battle. Experience affects the number and quality of solutions that they outline (Knighton 2004). Second, the four battalions that allied their forces in the Hong Kong battle have never fought together before and thus had no history of interaction. Third, cognitive biases were again at play, in decisions made during the battle. Fourth, bad luck or inadvertent mistakes, also played a role in the final outcome. Fifth, dishonesty played a role in determining behaviour, decisions and outcomes.

The rational decision-making model would predict that, having decided to defend rather than surrender, the colony would not be able to fight for more than a couple of days in the face of the mighty Japanese invasion. The seventeen days defence of Hong Kong was surely helped by good plain luck. About a third of the Japanese shells fired in the battle of Hong Kong were defective and did not explode. The Japanese were also unfamiliar with the terrain, and the winding, narrow roads made transportation difficult. Moreover, the Japanese had not cut army and civilian telephone lines, so these lines remained in use until almost the end of the war (Lindsay 2005). The naturalist theory based on a moral and ethical explanation of decision-making can also be used to explain the outcome of the battle. According to this theory, people’s behaviour is influenced by their ethical and moral beliefs, situated deeply at a bedrock level (Knighton 2004). Although it was only rational, perhaps, for the troops to surrender, their deep sense of patriotism and selflessness determined their behaviour in the last days of the battle. Tremendous acts of incredible courage and self-sacrifice were performed by the Allied Forces in Hong Kong. Company Sergeant Major J.R. Osborn who fought on Mount Butler rolled over onto a grenade and by this action saved the lives of several troops under his command. Hughesiliers from the Volunteers battalion, showed tremendous courage through their actions. Although the Hughesiliers were given the passive responsibility of preventing sabotage at an electrical plant, they chose to actively fight in the face of the Japanese attack and thus lost their lives (Lindsay 2005).

Decision to Surrender

According to the rational model of decision-making, the Allied Forces should have surrendered much earlier than they did. Continuing fighting would only lead to more casualties, without any benefits. However, emotions and traps of decision-making such as escalation of commitment, contributed to the decision to fight on. Four days after the withdrawal from the mainland, on

December 17th, Sir Mark Young, the Governor of Hong Kong, refused to surrender, despite reports from Brigadier Wallis and others that the island was not defensible with the outnumbered forces and arsenal available. Surrendering would damage his credibility and reduce the chances of a possible promotion (Lindsay 2005).

OUTCOME OF THE BATTLE

The eventual outcome of the Hong Kong battle is not unexpected. As Lai (1999) has noted, the fact that Hong Kong did fall to the Japanese belies the fact that the defenders had performed far better than anyone expected. What is unexpected, though, is that it took 17 days before the defenders surrendered. What is so remarkable about a battle that lasted a mere 17 days? We need to briefly look at another battle that took place scarcely two months later.

When World War II began in the Pacific, the British Empire boasted its Gibraltar of the East in the island fortress of Singapore. With a garrison of 80,000 troops, in addition to a further 50,000 further up in the Malay Peninsula, and supported by large calibre naval guns as well as capital ships, Singapore was thought to be an unsinkable fortress that would provide the anchoring point for Allied military campaigns against Japan. Unlike the resources (or lack thereof) allocated to the defence of Hong Kong, vastly better-trained and supplied personnel and equipment were made available to the Singapore fortress. In addition, the attacking Japanese troops were actually fewer in number (estimated at only 30,000) than the defenders. Yet only seven days after the Japanese landed on the island of Singapore, the Gibraltar of the East fell.

We argue, based on this brief comparison, that although the defence of Hong Kong was unsuccessful, it was a remarkable achievement nonetheless. Indeed, as we have already noted, Japan did not launch its attack on Singapore until February 1942. Certainly, an earlier surrender in Hong Kong would have allowed the Japanese army to launch an assault on Singapore even sooner. Albeit ultimately in vain, the prolonged defence of Hong Kong bought time for the British to reinforce Singapore for the inevitable assault. For these reasons, a study of the defence (and fall) of Hong Kong in 1941 provides valuable lessons in strategic decision-making, and the (arguably) unexpectedly long contest issues an interesting case study.

Factors that affected the three sets of decisions and ultimately determined the outcome of the battle are represented graphically in Figure 2. The side, rectangular, boxes indicate elements that were outside of the control of decision-makers involved in the battle, i.e., “external” elements. There was little Major General Maltby and his officers could do about good or bad luck, the physical terrain of the island, the lack of training and the resources that they were given. Whether they would fall in cognitive traps or not, external elements would still impact the outcome of the battle. Thus, the quality of decision-making was constrained by these elements, which existed independent of their will.

“Internal” factors are represented in the oval shapes. These are decision-making biases, emotions, risk taking propensity, problem framing. These are elements that, if controlled, would perhaps increase the quality of decision making at all three points.

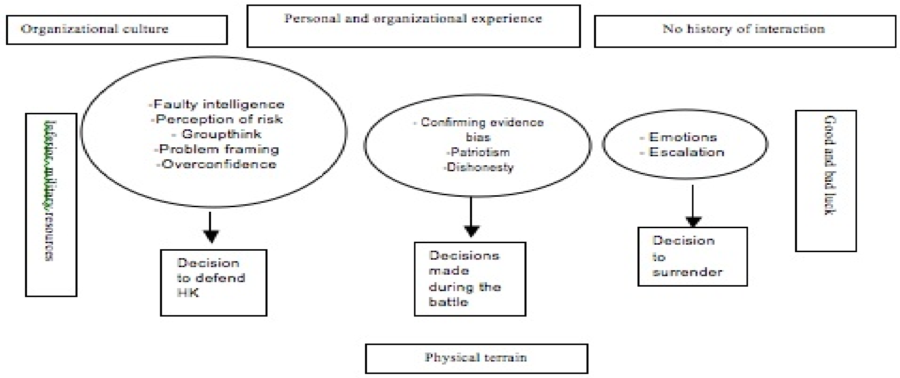


Figure 2. Factors affecting decisions and the final outcome. Rectangular boxes indicate 'external' factors, while oval shapes indicate 'internal' factors.

CONCLUSION

Clearly, the strategy of the Allied Forces in Hong Kong was not hard to figure out. It was defending Hong Kong with the available resources, as long as possible. However, execution of strategy was what ultimately rendered Hong Kong "hostage to fortune." Elements such as leader's personality and their perception of risk, emotions, overconfidence, cognitive biases, organizational culture and structure, all had an impact upon strategy implementation. Another common sense, traditional view, is that strategy is formulated at the top and executed by the lower organizational levels. This view also implies a separation of the leadership function of "doing the right thing" from management, defined as "doing things right." Consequently, the pragmatic function of "management" is held to a lower esteem than the more fascinating function of "leadership" (Mintzberg 2009). However, there is growing research on how these lower organizational levels have a tremendous contribution in fundamentally changing, formulating organizational strategy and sometimes even obstructing strategy formulated at the top (Beer and Eisenstat 2004; Bower and Gilbert 2007; Pfeffer and Sutton 2006).

High stakes decisions are more susceptible to cognitive biases and psychological traps because they are based on the most assumptions and require the involvement of many people (Hammond, Keeney and Raiffa 2006). The decision to defend Hong Kong in the face of the Japanese invasion, decisions made during the battle and the decision to surrender were all major, critical decisions, especially susceptible to such biases as overconfidence, problem framing, availability heuristics and confirming-evidence. Making decisions under uncertainty has been an important attribution of military commanders as well as CEOs. However, in contrast to decisions made by business leaders, military decision-making is a matter of life and death. Thus, military leaders have a greater responsibility in making accurate decisions (Knighton 2004). Different perceptions of risk, as well as personality characteristics, emotions and team dynamics, affect the process of decision-making, which is susceptible to many cognitive biases.

References available upon request from Ed Chung, chunge@etown.edu

CURRENT RESEARCH ON THE AFFECT OF ORGANIZATIONAL CHANGE ON TEAM DYNAMICS

Nicole Jean Christian, Walden University, (631) 732-8969, nicolejchristian@gmail.com

ABSTRACT

The focus of this paper is the affect of organizational change on work teams, specifically, the affect of organizational change on organizational culture during the organizational change process. Current theories of organizational change were analyzed based on a critical review of relevant studies. The findings of this critical review build the framework for application to a "real world" situation in the author's dissertation research. In this paper, the author will identify the key organizational factors affected by organizational change, specifically as they pertain to team dynamics. The desired result for this research will be to identify and synthesize the most effective processes to date to use for successful organizational change initiatives in contemporary organizations. These "best practices" are identified with a particular emphasis on improving team communication, interaction and dynamics during the change process. A particular emphasis in the author's dissertation research will center on intergenerational and cultural shifts at work and how these dynamics are altering the landscape of the 21st century workplace. This paper is the author's first in-depth review of the current literature about this particular aspect of organizational change.

INTRODUCTION

By all accounts the workplace is aging. The federal government and other established sectors of our economy have acknowledged the "aging" of the workplace. For the first time in history there are four different generations, "struggling to coexist" [11, p. 15] in the workplace. These four generations (Mature, Baby-boomer, Generation X and Generation Y) have distinct cultural perceptions and experiences. The mature generation is described as anyone born before 1946; the baby boomer generation are those born between 1946 and 1964; Generation X is anyone born between 1965 and 1980 and finally Generation Y are those born between 1980 and 1990 [11]. All of these different people share workspaces, sometimes serve on the same teams and often work for one another. Each of these generations has been raised with different levels of exposure to technology and collaborative work processes. When these groups work collaboratively sometimes, the use of different styles of communication and behavior can cause conflict. These generational differences in style, language and behavior are a catalyst for much of the organizational change efforts many of us experience today.

Intergenerational and cross-cultural communication at work is altering the landscape of the 21st century workplace. Along with different cultural values are the added factors of work motivation, employee engagement (involvement) and organizational culture. Different generations are motivated based on different cultural experiences and meaning. Also at play in the workplace is the specter of what stage of life (young, middle aged, mature) an employee is in during organizational change processes. What are organizations to do now that all of these issues are converging at once for the first time in history in the workplace? How can organizations prepare workers and organizational culture for such a shift in perspective where once the populations of most workplaces were so similar and are now so different? As the workplace continues to change based on these and other economic and demographic forces organizations can gain greatly from identifying those factors that influence team dynamics and what affect organizational change has on team dynamics. This paper will present and identify three key factors that influence team dynamics during organizational change processes: work motivation, employee engagement, and organizational culture.

FACTORS THAT INFLUENCE TEAM DYNAMICS DURING ORGANIZATIONAL CHANGE

Work motivation

Ackerman and Kinder (2004) completed an extensive longitudinal and cross sectional study on the effects of aging and development on work motivation [1]. They identify and explain what they refer to as the "four common themes based on the nature of change across the life span: loss, growth, reorganization and exchange" [1, p. 441]. These four stages in adult development influence employee work motivation. These four stages of adult development become a prism through which an individual sees life and their goals and dreams. Organizational change efforts will be influenced by the stage of adult development a person is in. Individual employee motivation and personal goals may change depending on what stage someone is at in their adult development. In order to ensure the highest level of employee work motivation and change acceptance, organizational change efforts must take these stages of adult development into consideration when implementing change.

Individual work motivation will be influenced by generational differences as well as individual age and time spent in the workplace. As it pertains to workplace teams, Athanasaw (2003) identifies four factors that influence effective team work behaviors: the number of years of professional experience, the number of teams an employee participated on, the type of training each member had and the number of courses, and finally how they became a part of the team in the first place [2]. All of these factors Athanasaw (2003) identifies affect how an individual reacts to change and team-based work [2]. Franklin Knowles (1998), considered the father of adult learning, explains that there is much potential to influence motivation in adults because "human behavior is goal oriented" [8, p. 189]; we are always striving to achieve and accomplish goals in our lives. It is important that organizational change efforts include specific efforts to enhance or influence individual motivation and goals. McGregor's [9, p. 45] theory of motivation is based on the premise that individuals want to do a good job and will do even better if and when their individual goals can be integrated with the organization's goals. He encourages managers to create an environment where individuals can simultaneously satisfy organizational and personal goals through their work behavior. In order to be successful, any change effort must actively involve the individual [10]. Involving people, valuing them and their efforts, can enhance their motivation and desire to work toward achieving organizational goals. If the goal is organizational change, then involving and valuing employee input during organizational change efforts is critical to achieving organizational success. Organizational change can and does affect individual motivation because change is influenced by individual values and goals and employee motivation is based on individual value systems.

Employee engagement

Employee engagement is a multi-faceted process. It is defined by authors Frank, Finnegan and Taylor (2004) as "how employees feel about their employer, its leaders, working conditions – and behavioral components – measures of intent to act certain ways, skills they choose to bring to bear, to go the 'extra mile' " [7, p. 15]. Engaged employees are more likely to stay with an organization and be active in its internal process and changes [7]. During organizational change it is important to emphasize those factors that will influence employees to stay and be engaged in the change process. Customizing rewards and incentives is one way to enhance retention and engagement during change. Trust is another factor. Trust between the employee and the employer will influence whether or not a worker will remain engaged in a change process [7]. Organizations must also create opportunities to build self-esteem and coach team members and leaders. When employees feel invested in they will value the organization and its change efforts. In order for team functions to become and continue to be effective, organizations must create climates that are enjoyable and fulfilling. They must also be flexible in addressing and identifying employee needs and views [7, p. 19-20].

Organizational change is no easy task. It is a difficult to require adults in the workplace to change the way they think, behave and communicate, especially when collaborative approaches are what most companies use. When people work together, differences abound. How can those differences create opportunity for success? Employees must see the change as important and necessary or organizations risk failure. Organizational change will influence employee engagement in that what employees value as an important change, "they will want to implement" [13, p. 3]. What employee's value they will support and what they collectively believe they will implement. Organizational change efforts must be rooted in a shared belief about what the change will accomplish and why it is necessary. When employees, "share a common, favorable assessment. . ." [13, p. 4] of a change process, and what it requires, they will be confident in its necessity and overall affect on the organization. A proper reward system can also enhance employee engagement. Change agents in an organization value being rewarded socially and monetarily [5]. Depending on the individual, their personality and value system, proper rewards can be developed that can positively influence change acceptance and ultimately employee engagement.

Organizational culture

Organizational culture is the shared beliefs, values and perceptions among individuals in an organization [4]. Organizational change has a powerful affect on organizational culture. The culture of an organization determines its readiness for change and flexibility. Organizations that are more ready to embrace change have a culture that encourages learning from the past and planning for the future [12]. Jim Collins (2001) points out that companies that make it from 'good to great' are companies that can embrace change and create an environment where "people have a tremendous opportunity to be heard and, ultimately, for the truth to be heard" [6, p. 88]. During organizational change processes the culture of an organization will go through constant shifts and transformation as its members create new meaning and behavior. During change, organizations must be poised to learn from the past and be prepared for whatever future challenges await. An organization's ability to learn and change will be a significant factor in determining its success and survival in an ever-changing market and economy [12]. Organizational change influences all aspects of organizational culture from the stories people tell each other to the design of employee orientation programs. In order for organizational change processes to be successful the change itself must be embraced by all members of the organization not just top management. Top managers may not control or direct change as once thought but rather middle managers and their staff have more influence than given credit for [3]. The way these individuals exercise some control over the change processes is through informal social networks. These social networks are pivotal in directing change processes and these social networks are often more the domain of middle managers and non-management level employees. The influence of social networks should be a part of the analysis as organizations begin to plan and implement change. If social networks are ignored, change participants can become disillusioned and frustrated with the process [3]. Informal relationships – social networks – are the fabric of organizational culture and this fabric is transformed when change initiatives are implemented. Organizations must have plans in place to lead a change process from the beginning until the end, including creating opportunities to discuss how the culture was when the change process began and how the culture will be different after the change process is complete.

CONCLUSION

Since societal and workplace needs are changing rapidly as a result of technological advances, changing demographics, globalization, economic uncertainty and crisis, the workplace has provided the basis for much of the corporate innovation we experience today. In order to adapt to the changes that technology, changing demographics and globalization have ushered in, today's manager and worker must be <u>flexible</u>, <u>adaptable</u> and <u>have the innate ability to accept and be comfortable with change</u>. There is a universal need for training and development in the area of organizational change. Organizations need the opportunity and space to discuss and identify where change can and will occur in various areas of their organizations. I successfully designed an eight-part 24 hour class for the Dowling Institute (Dowling College, Oakdale, NY) for its corporate training program. The course was held for eight weeks in the winter of 2009 and met for 3.0 hours per week. The nine participants were guided through a variety of change management workplace scenarios, group exercises, presentations and discussion. The course was designed to prepare participants to meet the needs of a rapidly changing work environment. Application and discussion of new ethical standards, diversity, teamwork and the sharing of knowledge/learning was covered in this course. The objective of the course was to provide the participant with the tools to <u>identify</u>, <u>address</u> and <u>facilitate change</u> among workplace teams, individuals and in the world. At the end of the course, participants developed a personal change management style to use at work and in their personal life. The long-term outcomes of this course and next steps will be the subject of a case study that I will write later this year.

Commitment to and understanding organizational change and its affect on team dynamics is a major undertaking for any organization. The current literature indicates that organizations that undergo organizational change can be more successful if they understand that change alters the core elements of an organization: motivation, engagement and culture. Offering training and development in these areas of change is a good place to start but more is needed if an organization desires success. The workplace itself must change – long term – and those changes must be supported by the entire organization. The organizational culture, language, imagery and team processes must change to reflect the shifts in our society. That means that the workplace of today with - for the first time - four different generations in it, is at the forefront of a major cultural paradigm shift at work. These individuals must see the benefits of each generations learned behavior and lessons and be able to collectively create a cohesive vision for the future.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the support and encouragement of Dr. Elana Zolfo, Vice President of Corporate Affairs (and Past-Dean of the Townsend School of Business), Dowling College; Dr. Maureen Mackenzie, Associate Professor of Management and Leadership, Dowling College and Assistant Dean Antonia Losciavo, Dowling College; Dr. Carol Wells, Professor of Business, Walden University; my husband Majai Q. Goodine, my mother Barbara Jean Christian, MBA, alumnus of Dowling College; the Christian family, and in memory of my father, Philip A. Christian.

REFERENCES

[1] Ackerman, P.L., Kanfer, R. "Aging, adult development, and work motivation," *Academy of Management Review*, 2004, 29(3), 440-458.

[2] Athanasaw, Y. "Team characteristics and team member knowledge, skills, and ability relationships to the effectiveness of cross-functional teams in the public sector", *International Journal of Public Administration*, 2003, 26(10/11), 1165-1204.

[3] Balogun, J., Johnson, G. "Organizational restructuring and middle manager sensemaking", *Academy of Management Journal*, 2004, 47(4), 523-549.

[4] Boan, D."Cognitive-behavior modification and organizational culture", *Consulting Psychology Journal*: *Practice and Research*, 2006, 58(1), pp. 51-61.

[5] Chrusciel, D. "What motivates the significant/strategic change champion(s)?" *Journal of Organizational Change Management, 2008*, 21(2), pp. 148-160.

[6] Collins, J., *Good to great: why some companies make the leap. . . and others don't*. New York, NY.: HarperCollins Publishers, Inc., 2001.

[7] Frank, F.D., Finnegan, R.P., Taylor, C.R. "The race for talent: retaining and engaging workers in the 21st Century", *Human Resource Planning*, 2004, 27(3), 12-26.

[8] Knowles, M., *The adult learner: the definitive classic in adult education and human resource development*. Woburn, Ma.:Butterworth-Heinemann, 1998.

[9] McGregor, D., *The human side of enterprise*. New York: McGraw-Hill Book Company, Inc., 1960.

[10] O'Toole, G. "Count me in: the role of action learning in making learning and skills provision more inclusive", *Action Learning: Research and Practice*, 2007, 4(1), pp. 69-75.

[11] Pareso, B., Piccoli, K. “Evolutionary road: the millennials, the youngest offspring of the baby boomers are entering the workforce. Is the world ready for them? Are they ready for the world?”, *Long Island Press*, October 22 – October 28, 2009.

[12] Reed, J., Vakola, M. “What role can a training needs analysis play in organizational change?”, *Journal of Organizational Change Management*, 2006, 19(3), pp. 393-407.

[13] Weiner, B. “A theory of organizational readiness for change”, *Implementation Science*, 2009, 4(67), doi: 10.1186/1748-5908-4-67 also available at (http://www.implementationscience.com/content/4/1/67).

SOME CONSIDERATIONS FOR THE USE OF STRATEGIC PLANNING MODELS

Robert F. Everett, 978-837-5406, Robert.Everett@merrimack.edu
Catherine Rich Duval, 978-837-5410, Catherine.RichDuval@merrimack.edu
Girard School of Business and International Commerce
Merrimack College
315 Turnpike Street
North Andover, MA 01845

ABSTRACT

Analytical models for strategic planning, particularly SWOT Analysis, have become an integral part of our business culture in both practice and academia. Yet, despite their widespread use, thoughtful scholars have, from time to time, criticized their use and proposed alternative methods for strategic analysis or, in some cases, proposed abandoning the use of strategic models altogether. Focusing on SWOT Analysis, this paper reviews some of the most relevant problems that have been discussed about SWOT in the context of how this tool is most commonly used. We then propose how SWOT can be used in new, more powerful ways.

Keywords: SWOT, Analysis, Strategy

A Brief History of SWOT Analysis

The origins of SWOT Analysis can be traced back to a 1965 text by Learned, et al [8] in which they described the basis of business strategy formulation in terms of matching organizational competence and resources to environmental opportunities and risks. From that point, the use of SWOT as a strategic tool has grown to the point where a Google search of term "SWOT Analysis" produces over 1.3 million hits, and a keyword search on Business Source Premier produces over 13,500 hits.

The Apparent Trouble with SWOT

Despite its clear popularity, SWOT has had its detractors. Wharton professor J. Scott Armstrong [1] advises "Don't do SWOT." Hill and Westbrook [4] argue that "It's Time for a Product Recall." Several other authors have made meaningful critiques of the tool. [2] [3] [5] [6] [7] [9] [12]

Valentin [12] (another opponent of SWOT) identified the following problems with SWOT Analysis:

- It yields banal or misleading results
- It has a weak theoretical basis
- It implies that organizational factors can be "neatly" categorized as positive or negative
- It encourages "superficial scanning and impromptu categorization"
- It promotes list building as opposed to thoughtful consideration
- It does not look at tradeoff among factors

- It promotes muddles conceptualizations, particularly between accomplishments and strengths

In general, we agree with these points. However, we propose that the problem is not with the fundamental concept of SWOT Analysis, but rather with how SWOT has come to be used in business and taught in the classroom.

The Real Trouble with SWOT

We believe that there are two fundamental problems with the common use of SWOT Analysis and it is these two problems that produce the deficiencies Valentin and others articulate.

The first of these problems is a simple overestimation of the power of SWOT as a tool for strategic analysis. SWOT Analysis should not be considered sufficient to produce the information and insights that a responsible strategic planner needs. Replacing SWOT with another single tool deemed to be better (e..g., [4]) may, in some circumstances, marginally improve results. However, it leaves the planner still dependent on a single tool. We use multiple tools to do many common tasks such as yard care, replacing a faucet washer, or styling hair. Why would we think that there has to be one, unique tool that satisfies all of our basic strategic analysis needs?

Another factor in the overestimation of the power of SWOT lies in the fact that it is usually done as a one-time task, rather than an ongoing process. Strengths and weakness must be identified evaluated in the context of specific perceived opportunities and threats. Opportunities and threats must be evaluated in terms of specific strengths and weaknesses. This mutual adjustment requires an iterative process, not a simple enumeration.

The second major problem relates to how a SWOT Analysis is performed and the expectations we have from it. As practiced and taught today, a SWOT Analysis usually is done as follows:

1. The individual or group performing the SWOT Analysis sets aside some time and place to address this task. The time may be as little as a few minutes or hours for an individual to a couple of days for a group to meet offsite to a few weeks for a designated committee.
2. Facts, beliefs, impressions, and ideas regarding organizational strengths and weaknesses are enumerated along with the potential opportunities and threats presented by the organizations environment.
3. These SWOT factors are then assessed by the group and consensus is sought regarding the importance of the factors identified.
4. Factors may then be weighted in terms of importance.
5. The individual then examines the organizations strengths and weaknesses in the context of environmental opportunities and threats.
6. Organizational strategies emerge from this examination.

The major reason why these issues are problematic for SWOT Analysis lies in the assumption that a SWOT Analysis, when comprised entirely of internally generated data, is even capable of providing an accurate and unbiased picture of an organization and its environment. As intuitively appealing as the SWOT process may be, there are some critical underlying assumptions regarding how it is most commonly done that are problematic at best.

First, there is the assumption that the information needed to assess the various SWOT factors actually resides within the individual, the committee, or even the organization performing the analysis. The experience of the authors, as consultants, has shown this to be highly unlikely. Both individuals and organizations have biases, preconceptions, misunderstandings, and gaps in their knowledge about themselves and their world. We have found, however, that it is these biases and gaps that may be the most important things for strategic planners to understand, because these biases may be creating the organizational problems the strategic planning effort is trying to correct.

Second, individuals within an organization may not even know the right questions to ask regarding themselves and their environment. There may even be the tendency to focus on certain issues in a SWOT Analysis not because those are the most important issues, but because those issues are most fully understood.

Third, even if the right questions were asked and the information was available within the organization, it would be difficult for that information to be accessible to any committee or individual performing a SWOT Analysis. There are a number of reasons for this:

1. Turf: Individuals like to think that they are experts in their own field of operations and are resistant to contradictory information from people in other areas.
2. Insecurity: Some people may not want to appear ignorant about areas where they are supposed to be knowledgeable.
3. Fear: Some people may be reluctant to contradict superiors when those superiors are wrong for fear of some sort of retribution.
4. Conventional wisdom: There may be some ideas so firmly ingrained in the organizational culture (e.g., “We produce quality products.” Or “We care about our employees.”), that those ideas are presumed to be true without challenge.
5. Ego: Some people in positions of authority may simply declare things to be true because they think so.
6. Expedience: Because a SWOT Analysis is usually a task people must do in addition to their main responsibilities, some people may be reluctant to fully engage difficult or controversial issues because doing so would take too long.

Finally, there may be internal resistance to the direction that a given SWOT Analysis is taking when that direction impinges on the performance evaluation of management authorities within the organization; or the individual or members of the committee performing the analysis.

The key point we are making here is that it is exceedingly difficult to perform a SWOT Analysis using internal information sources and get to the truth, without substantial bias.

The Real Benefits of SWOT

This does not mean that an internally done SWOT is not an exceedingly valuable exercise. It simply produces valuable results other than the ones for which most managers look. We believe that there are at least three of these additional benefits.

1. Assessing organizational belief systems: An internally based SWOT Analysis can also help an organization assess its own belief system. If we accept that SWOT Analysis, as commonly done, may contain a wide variety of biases, preconceptions, and other deviations from objective accuracy about the “real” nature of the organization and its

world, it may, in fact, present a quite accurate view of organizations internal beliefs about itself and its world.

By this suggestion, we mean that, to an astute, unbiased observer, the inconsistencies, inaccuracies, omissions, and distortions contained in an internally done SWOT can reveal a great deal about how organizational decisions are made. In particular, one can examine:

- The extent to which strengths may be overestimated and/or weaknesses underestimated
- Inconsistencies in the evaluation of certain factors and which individuals within the organization hold the contradictory views
- The extent to which those involved in the SWOT process are willing to think out of the box or support those that do
- The presence of systemic or individual preconceptions or biases regarding customers, markets, and competitors.
- The extent to which identified preconceptions or biases are defended when challenged.
- The extent to which a searching and fearless examination of truth is valued in the organization.

2. Generating questions: As any trained researcher will attest, the value of any answer depends on the question being asked. When the focus of attention is on identifying the right questions to ask instead of attempting to identify "truth," a SWOT Analysis can become a very stimulating form of brainstorming.

   As issues are identified, the following questions can be raised for discussion:

   - What current beliefs do we have regarding this question?
   - For each belief, how do we know this? By assumption, judgment, or evidence?
   - Is there someone within our organization that knows more about this issue than I/we do?
   - How would I go about gathering evidence to clarify or test this belief?
   - How would I classify this issue?
     - An accurate assessment is critical for strategic success
     - An accurate assessment would be helpful
     - An accurate assessment would be nice, but not worth investing in

3. Identifying research agendas: Once potential areas for inquiry have been identified and their importance assessed, a program for more formal organizational assessment and/or market research can be formulated. This agenda would also have to take into account timelines and available resources to invest in the effort. However, given that the most critical research issues would have been identified and there would be clear justifications for the investments made, it would be more likely that top management would invest in these, rather than ad hoc, research programs.

4. Providing inputs to other analytical processes: By using a SWOT Analysis to generate issues, questions, and awareness of potential biases, other analytical methodologies (e.g., BCG Analysis, GE Business Screen Analysis, SPACE Analysis [10]) can be fed more well-considered and vetted raw material.

5. Helping to understand the complex and dynamic nature of an organizations operating environment. By examining how perceptions of strengths and weaknesses vary in the context of various identified opportunities and threats, management can develop a greater appreciation for the truth about its own organization and be better able to prioritize its own organizational development investments.
6. Part of an ongoing strategic planning system: Finally, we would recommend making strategic analysis in general and SWOT Analysis in particular, part of a continual assessment program. Organizations monitor financial performance on a continual basis and invest substantially in executives, personnel, and computer system to do so. In the same way, organizations should track internal and external environmental issues with at least the same diligence. After all, that is where the dollars come from.

References

[1] Armstrong, J. Scott (2009) "Don't do SWOT, A Note on Marketing Planning" Manyworlds.com

[2] Bernroider, Edward (2002) "Factors in SWOT Analysis Applied to Micro, Small-to-Medium, and Large Software Enterprises: An Austrian Study," *European Management Journal*; Vol. 20 Issue 5, p562.

[3] Chermack, Thomas J. and Bernadette K. Kassahanna (2007) "The Use and Misuse of SWOT Analysis and Implications for HRD Professionals," *Human Resource Development International*; Vol. 10 Issue 4, p383-399.

[4] Hill, Terry and Roy Westerbrook (1997) "SWOT Analysis: It's Time for a Product Recall," *Long Range Planning*; Vol. 30 Issue 1, p46-52.

[5] Edwards, Colin C. (1993) "What's SWOT?" *Management Services*; Vol. 37 Issue 6, p38-38.

[6] Houben, G. and K. Lenie (1999) "A knowledge-based SWOT-analysis system as an instrument for strategic planning in small and medium sized enterprises,".*Decision Support Systems*, Vol. 26 Issue 2, p125.

[7] Humphreys, John (2007) "Weakness or Opportunity?" *MIT Sloan Management Review*; Vol. 48 Issue 3, p96-96

[8] Learned E. P., C.R. Christensen, K.E. Andrews and W.D. Guth, *Business Policy: Text and Cases*, Irwin, Homewood, IL, 1965.

[9] Lee, Kuo-liang and Shu-chen Lin (2008) "A fuzzy quantified SWOT procedure for environmental evaluation of an international distribution center," *Information Sciences*; Vol. 178 Issue 2, p531-549.

[10] Rudder, Luetitia and Lynette Louw (1998) "The SPACE Matrix: A Tool for Calibrating Competition," *Long Range Planning*, Vol. 31 Issue 4, p549-559.

[11] Valentin, Erhard K. (2005) "Away With SWOT Analysis: Use Defensive/Offensive Evaluation Instead," *Journal of Applied Business Research*; Vol. 21 Issue 2, p91-104.

[12] Yüksel, İhsan and Metin Dagˇdeviren (2007) “Using the analytic network process (ANP) in a SWOT analysis – A case study for a textile firm,” *Information Sciences*; Vol. 177 Issue 16, p3364-3382.

LEARNING REAL WORLD EXPERIENCE IN THE CLASSROOM

WORKSHOP AUTHOR
Christopher Roethlein, Bryant University, 1150 Douglas Pike, Smithfield, RI 02917
(410) 232-6721
croethle@bryant.edu

INTRODUCTION

Many undergraduate students have never been inside a manufacturing facility and lack service related experience. This lack of experience creates a situation where students are often unable to relate to some operations management textbook and classroom discussion material. Furthermore, what often appears to be very simple in the text is much more complicated in the real world. This workshop presents a semester based learning experience that gives the students experience in quality concepts, lean techniques, communication with internal and external customers, technical writing and value added services.

The Production Game

The Production Game (Denton, 1990) is designed to simulate real life production situations. The class is divided into individual companies and the objective of each company is to try to develop an effective and productive organization that is competitive in the marketplace. *The Production Game* provides lessons in communicating and understanding customer's definitions of quality, teamwork and an appreciation of how quality requirements are communicated within a supply chain. Within four or five production days, students learn how to better communicate with their customer and fellow team members. They also learn how to expedite production efforts and build a company strategy around the team's competitive strengths. Financial, efficiency and productivity metrics are established and monitored throughout the game. The principles of lean techniques that increase efficiency, remove bottlenecks and improve overall workforce efficiency are explored. A formal, technical report is required at each step of the game and a final, culminating report demonstrates their mastery of the concepts. *The Production Game* is a fresh alternative to lecturing that gets the students involved and realizing how they can implement the OM techniques discussed in class.

Abstract

This paper adds to the ongoing development of the field of quality management research. The purpose of this paper is to clarify the status of quality management research and categorize the recent literature in quality management and to better position the field for growth and maturity. We performed a rather extensive overview, review, and discussion of quality-related literature since 1995. We used centering resonance analysis (CRA), which has not been previously utilized in the quality management literature, to analyze 1877 research articles to identify influential word pairings in the literature overall, and in 172 articles appearing in recognized “A” journals. These provide a framework for reviewing the literature to categorize and summarize articles and key research contributions, identify trends, recommend areas for future research in the field of quality management.

OPERATIONALIZING SERVICE QUALITY: PROVIDERS' PERSPECTIVE

Linda Boardman Liu, Simmons College, 617-521-2412, linda.liu@simmons.edu

ABSTRACT

The focus of this study is on understanding the ways that technical and functional quality are operationalized and measured in a typical service encounter such as a customer would experience in calling a bank, cable provider, or other call-center environment. Using a two-part panel study with experts in the call center field, we find that organizations are focused on providing service that is *accurate* and *fast*. In addition, we identified a set of activities organizations use to ensure service quality: employee development, one call resolution, effective technical infrastructure, organizational culture, process speed, and answer quickly.

Keywords: quality, service operations

OPERATIONALIZING SERVICE QUALITY

Quality is key to any business, and is one of the critical areas where a company can gain competitive advantage. But what exactly is quality? For goods, quality is a well-defined construct, consisting of readily measurable elements such as conformance to specifications, absence of defects, or performance as expected [5]. For services, there are many approaches to defining quality, which complicates both theoretical and empirical research on service quality. There is, however, agreement that quality is important as the link between service quality and customer satisfaction is clear [2] [1] [7]. Despite the large body of research on service quality, it remains difficult to define, measure, and manage service quality.

Designing a service experience that creates this overall perception of service quality is a challenge. Realistically, a provider will design a system that attempts to control technical quality, as these aspects of service quality are more readily defined and institutionalized than functional quality, which is more closely connected to the interactive experience between the customer and the actual service provider. For providers, this is the difficulty: how are these aspects of service quality, the technical and functional dimensions [6], related to each other and related to the ultimate assessment of service quality: customer perception of service quality.

Research Setting

This research is set in the call center environment, a mass service process type by Schmenner's classification [9]. Call centers are deployed throughout the world as a cost-effective way of enabling customer-company interactions. A call center provides front-line contact to customers. An inbound call center is generally accessed by a customer initiating a call to the company. The call is routed to a group of agents usually with the help of a call distribution platform and/or voice response unit in order to route the caller to the 'right' agent, depending on the firm's choice of segmentation.

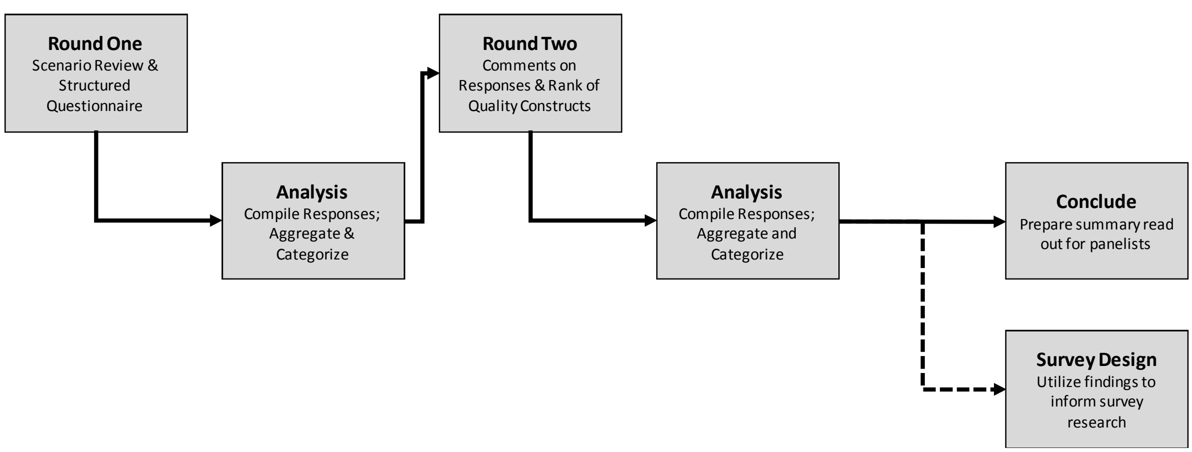
Panel Study

The focus of the study is on understanding the ways that technical and functional quality are operationalized and measured in a typical service encounter such as a customer would experience in calling a bank, cable provider, or other call-center environment. In order to better define the technical and functional quality constructs, we conducted a two-part panel study. A panel study enables the researcher to access experience and expertise in the research. By design, panels are conducted so the participants are anonymous to each other, in order to reduce any social effects that could result from the interactions among the panelists. Panel research is "useful in defining terms" [4] and produces an expert consensus or agreement. Panel research is qualitative in nature, and by design yields rich written descriptive information on the topic under study, grounded in realism.

The panel study included participants who are practitioners in the service operations field, with a particular expertise in the mass service environment, specifically call centers. Panelists were drawn from various industries and have roles that are both customer-facing and firm-facing. The entire panel study was conducted via email, and took approximately eight weeks to complete. The participants were a mix of both practitioners and consultants in the contact center arena. Seven of the participants were practitioners: call center directors or vice president of operations. Six of the participants were consultants, with roles that included principals, senior consultants, and consultants. On average, the participants had worked in or managed a contact/call center environment for 17.5 years (median = 15 years, min = 4 years, max = 35 years) and represented a wide variety of industries including banking, information technology, communications, insurance, finance, utilities, health care, and hospitality.

The panel participants were engaged in a two-round process. In Round 1, we captured broad concepts and ideas about service quality; Round 2 narrowed the focus to identify and rank specific measures of technical and functional quality (Figure 1). For Round 1, we distributed 21 questionnaires, and received 13 responses within two weeks. For Round 2, we distributed 13 questionnaires and received 11 responses within two weeks. These results provide insight into the operationalized definition and measurement of technical quality, functional quality, and overall service quality.

Figure 1 – Panel Study Design



Findings

In the first round, we described a generic contact scenario to ground the participants in a typical call encounter. Our goal was to gather descriptions of the kinds of *experiences* a customer will have in a quality service encounter as well as identify specific *activities* that occur in the contact center related to service quality. We used open coding to analyze and reduce the written responses. In the second part of the panel study, we used the responses to the first panel survey to refine and clarify the types of experiences, processes and measures associated with service quality.

We presented the 13 typical call center measures from Feinberg, et al [3] in alphabetical order. The participants used a Likert scale to indentify the importance of these measures to service quality, where 1 = very unimportant and 7 = very important. First Call Resolution (FCR), described as the percentage of customers who have satisfactory resolution on the first call, was identified as most important relative to service quality (mean=6.846). The second highest measure (mean = 6.077) was Average Speed of Answer (ASA), the average amount of time a customer spends waiting for the agent to answer (Table 1).

Table 1 – Call Center Measures

Measure	Mean	Standard Deviation
First Call Resolution	6.846	0.376
Average Speed of Answer	6.077	0.641
Blocked Calls	5.769	1.092
Employee Churn Rate	5.615	1.044
Service Level	5.615	1.121
Abandon Rate	5.538	0.967
Adherence to Schedule	5.000	1.732
Abandon Time	4.769	1.739
Average Handle Time	4.385	1.325
Average Work Time	3.615	1.660
Maximum Call Duration	3.615	1.121
Calls / Agent	3.231	1.536
Total Calls	2.923	1.706

Panelists were provided an alphabetical list of the customer experiences that were identified as being part of a quality service encounter. We asked the panelists to evaluate the *importance* of these experiences to service quality using a 7-point Likert scale, with 1 indicating “very unimportant” and 7 indicating “very important.” Table 2 presents these findings ordered from most important to least important.

Table 2 – Importance of Call Center Experiences

Experience	Description	Mean
Knowledgeable Agent	Agent has all the information needed…able to answer… knowledgeable…information available	6.909
Agent Demeanor	Easy to understand…polite, professional… courteous…friendly	6.273
One Call	Issue addressed in one call…resolved the first time	6.091
Good Technology	Technology screens and supports calls… system issues don't slow down the process	5.909
Limited Transfers	Customer not transferred…resolved with possibly one transfer…if necessary, a warm transfer	5.636
Rapid answer	Call should be answered quickly…promptly… reasonable amount of time	5.636

The panelists were presented with an alphabetical list of specific activities that occur in a contact center that were identified as having a significant influence on service quality. They were asked to read each statement carefully and use the scale to indicate how important each activity is to service quality, with 1 indicating "very unimportant" and 7 indicating "very important." Table 3 presents these findings ordered from most important to least important.

Table 3 – Importance of Call Center Activities

Activities	Description	Mean
Employee development	Effective and comprehensive…initial and ongoing training… coaching	6.300
Culture	Empathy…agent puts personal stamp on encounter… attitude… agent fits overall culture	6.100
One call	First call resolution…issue answered the first time	6.100
Technical support	Knowledge management systems…appropriate technology… technical access via internet or VRU…technology captures and uses customer data	6.100
Process speed	Quickly determine needs and resolve quickly…best possible service with least time investment on part of customer	6.000
Answer quickly	Speed to answer… wait time…answer promptly	5.700

Using the same list of activities that occur in a contact center, the panelists were asked to indicate how widely adopted these activities are in a typical contact center environment. The panelists were asked to think about a call center environment that they were most familiar with and to use a 7-point Likert scale to indicate if the identified activities were adopted in this call center environment, with 1 indicating "never" and 7 indicating "always." Table 4 presents these activities ordered from most frequently adopted to least adopted.

Table 4 – Adoption of Call Center Activities

Activities	Description	Mean
Answer quickly	Speed to answer… wait time…answer promptly	5.909
Process speed	Quickly determine needs and resolve quickly…best possible service with least time investment on part of customer	5.364
One call	First call resolution…issue answered the first time	5.000
Culture	Empathy…agent puts personal stamp on encounter… attitude…agent fits overall culture	4.727
Technical support	Knowledge management systems…appropriate technology… technical access via internet or VRU…technology captures and uses customer data	4.727
Employee development	Effective and comprehensive…initial and ongoing training… coaching	4.636

Finally, we asked the panelists to rank the importance of various competitive positions in the overall management of a contact center. We presented the respondents with an alphabetical list of typical competitive priorities that may be adopted by an organization and used to differentiate itself from its competitors. We asked the panelists to rank order the importance of each competitive priority for a contact center environment, with 1 indicating "the most important" and 5 indicating "the least important." Table 5 lists the competitive priorities from most important to least important.

Table 5 – Call Center Competitive Priorities

Competitive Priority	Rank
Service: value-added service, customer satisfaction	1.636
Quality: high quality, effectiveness focused	1.909
Responsiveness: fast and consistent response time	3.273
Cost: low cost, efficiency focused	3.818
Flexibility: variety of choices, ability to respond to change quickly	4.091

Discussion

We see that organizations are focused on providing service that is *accurate*, as demonstrated by the emphasis on first call resolution and low/no transfers and *fast*, by measuring waiting time (through abandon rates, service levels, abandon times, and speed of answer) and process time (through call handle times and agent work times). Panelists report that customers should experience encounters with skilled, professional, empathetic agents. The most important activities identified to create these contact center environments are employee development, one call resolution, effective technical infrastructure, organizational culture, process speed, and answer quickly. When the level of adoption of these activities is considered, we find that the most important activity, employee development, is the least adopted activity. Similarly, the least rated activities, process speed and answer quickly, are the most adopted activity. This apparent disconnect between what the providers believe is important to customer quality and what is emphasized in management practice demonstrates the Parasuraman, Zeithaml & Berry Gap Model [8] in practice: although practitioners have clearly formed beliefs about what is important to customers in terms of quality, they do not always manage their service delivery system in accordance with those expectations. In addition, the call center measures that are identified as most important also emphasize this focus on process: first call resolution, service level, speed of answer and managing waiting times are identified as the most important measures. Finally, as our data shows that these contact center environments are positioned to deliver value-added service, customer satisfaction and a high quality experience, this disconnect between what is believed to be important to customers and what the organization emphasizes in its operations becomes a critical issue in competing effectively.

References

[1] Aga, M. and O. V. Safakli (2007). "An Empirical Investigation of Service Quality and Customer Satisfaction in Professional Accounting Firms: Evidence from North Cyprus." Problems and Perspectives in Management 5(3): 84 - 98.

[2] Chen, C.-N. and S.-C. Ting (2002). "A Study Using the Grey System Theory to Evaluate the Importance of Various Quality Factors." The International Journal of Quality and Reliability Management 19(6/7): 838-860.

[3] Feinberg, R. A., L. Hokama, et al. (2002). "Operational Determinants of Caller Satisfaction in the Banking/Financial Services Call Center." The International Journal of Bank Marketing 20(4/5): 174-180.

[4] Flynn, B. B., S. Sakakibara, et al. (1990). "Empirical Research Methods in Operations Management." Journal of Operations Management 9(2): 250.

[5] Garvin, D. A. (1987). "Competing on the Eight Dimensions of Quality." Harvard Business Review November/December: 101 - 109.

[6] Grönroos, C. (1988). "Service Quality: The Six Criteria of Good Perceived Service Quality." Review of Business 9(3): 10-13.

[7] Hume, M., G. Sullivan Mort, et al. (2006). "Understanding Service Experience in Non-profit Performing Arts: Implications for Operations and Service Management." Journal of Operations Management 24: 304-324.

[8] Parasuraman, A., V. A. Zeithaml, et al. (1985). "A Conceptual Model of Service Quality and its Implications for Future Research." Journal of Marketing 49(Fall 1985): 41 - 50.

[9] Schmenner, R. (1986). "How Can Service Businesses Survive and Prosper?" Sloan Management Review 27(3): 21-33.

IN QUALITY MANAGEMENT, FIRM SIZE MATTERS

S. Bruce Han, Girard School of Business and International Commerce, Merrimack College, North Andover, MA 01845, (978) 837-5319, Bruce.Han@merrimack.edu
Maling Ebrahimpour, College of Business, University of South Florida - St. Petersburg, St. Petersburg, FL 33701, (727) 873-4154, bizdean@stpete.usf.edu

INTRODUCTION

The demand for ISO 9001 registration is increasing rapidly, as many companies require that their suppliers obtain it. It is also believed that attaining ISO certification is beneficial overall, helping organizations to achieve improved operational performance. In this study, we would like to investigate whether ISO 9001 certification efforts lead to TQM practices and improvement in firm's performances.

The following research questions will be empirically investigated in this study.
1. Do ISO 9001 certification efforts have a significantly positive effect on TQM practices?
2. Do ISO 9001 certification efforts have a significantly positive effect on firm performance?
3. Do TQM practices have a significantly positive effect on firm performance?
   Do size of a company matter in terms of answering the above three questions?

LITERATURE REVIEW

Today, more than 400,000 firms are certified worldwide with the ISO standards (Foster, 2007). Despite its overwhelming popularity, there is still considerable confusion surrounding the effects of ISO 9000 certification efforts on the TQM practices and firm performance (Sila and Ebrahimpour, 2003).

Some claim that obtaining ISO 9001 certification is the first step towards total quality and it can be a meaningful component of TQM (Askey and Dale, 1994; Porter and Tanner, 1996; Frehr, 1997; Anderson et al., 1999; Gotzamani et al., 2006). Anderson et al. (1999) contend that firms seek ISO 9001 certification as a way of achieving competitive advantage through quality management. Douglas et al. (2003) indicate that ISO 9001 certification leads to improvement in quality. Naveh ans Marcus (2005) found that installing and using ISO 9001 lead to achieving competitive advantage through improved on-time delivery and reduction in cost. Han and Chen (2007) suggest that ISO 9001 registration efforts enhance quality, cost reduction, dependability, and flexibility.

Some claim that ISO 9001 certification has little relation to TQM and is a bureaucratic procedure for international trade (Brecka, 1994; Stratton, 1994). Askey and Dale (1994) have found that firms tend to revert to their traditional practices after gaining certification rather than to move forward towards total quality management. McAdam and McKeown (1999) found that many small firms in Northern Ireland are benefiting from ISO 9001; however, the majority of them are not making progress towards TQM. Poksinska et al. (2006) also found that small organizations implemented ISO 9001:2000 with minimum effort and thus little change and benefits was achieved, losing many opportunities for improvement. Martinez-Loente and Martinez-Costa (2004) suggest that some ISO 9001 principles and TQM philosophy contradict each other and implementing both TQM and ISO 9001 simultaneously are not beneficial to firm's operating performance. Dreyfus et al. (2004) also indicated that a firm's ability to implement effective TQM is only marginally enhanced by ISO 9001 efforts.

Lakhal et al. (2006) revealed that there is a positive relationship between TQM practices and organizational performance. There are many studies that link the TQM approach and superior financial performance (Porter and Tanner, 1996; Escrig-Tena, 2004). Choi and Eboch (1998) found that TQM practices have a strong impact on firm's performances. Narasimhan and Jayaram (1998) found that TQM practices lead to competitiveness through improvement in competitive dimensions such as quality and cost.

The research objective is to conduct an empirical study based on real data to better understand the relationships between ISO 9001 certification efforts and TQM practices, as well as their effects on firm performance. Thus, a model is developed and proposed based on the literature (see Figure 1). Finally, multiple group analyses are performed on the model by splitting the data set by the firm size (Small vs Large) in terms of the number of employees.

METHODS

The survey instrument utilized a five-point Likert scale with 1 representing the low end of the scale and 5 representing the high end.

The population for this study consisted of ISO 9001 registered manufacturing firms operating in the United States with Standard Industrial Classification code 3600. A total of 1600 firms that belonged to the electronic industry were surveyed. The survey forms were sent to the ISO 9001 champions of these firms. Of the 1600 mailed surveys, 295 usable surveys were returned giving an 18% response rate.

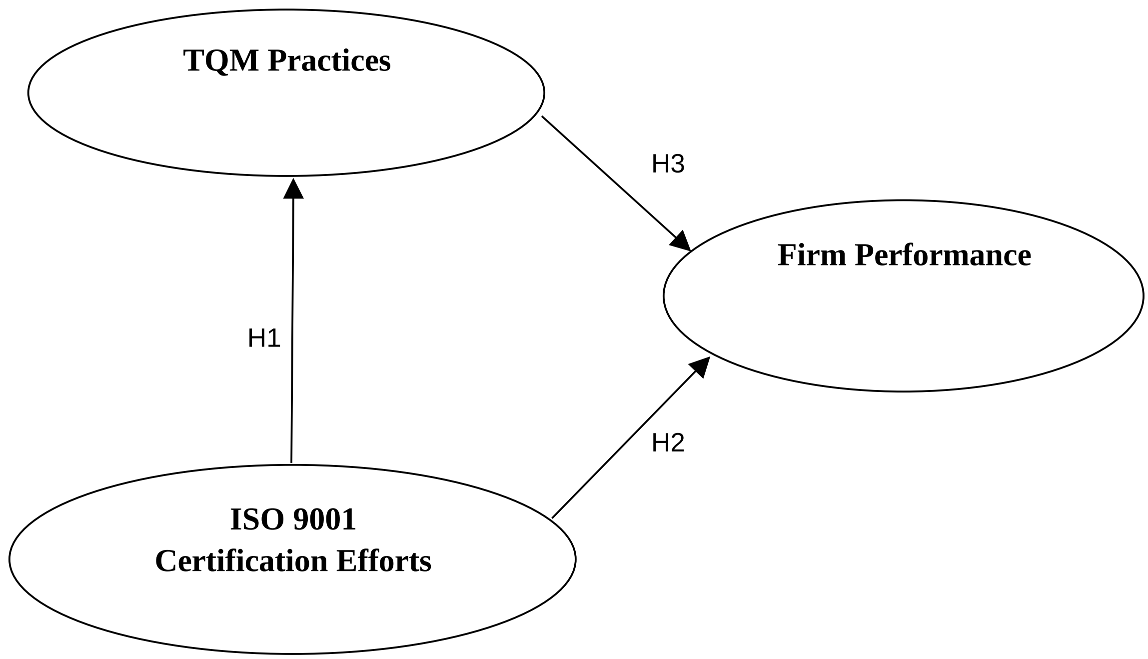
RESULTS

Table 1 presents the reliability and convergent validity of the constructs. All indicators showed significant loading at the .001 levels which provide evidence for convergent validity (Bagozzi and Yi, 1988). Crobach's Coefficient Alpha was used to assess reliability. All alpha values greater than .70 addresses the reliability of the constructs (Nunnally, 1967). Table 2 shows that all constructs are positively correlated and they are well below the suggested cutoff point of .90 (Bagozzi et al, 1991). This provides an evidence for discriminant validity.

MULTIPLE GROUP ANALYSES

The findings are summarized in Table 3. As can be seen from Table 3, all three models indicate adequate model fit.

Figure 1: Conceptual Model



Two hypotheses (H2 and H3) are consistently supported at the .01 significance level across firm size. The comparisons show that firm size has a significant effect on the hypothesis H1 that is being tested. H1 is supported by firms belong to the large size firms. In this study, large size firms are firms with one hundred or more employees and small size firms are firms with less than one hundred employees (Pagell & Halperin, 1998). H1 is not, however, supported by the small size firms.

DISCUSSION

This study was developed from a theoretical foundation and provides a deeper insight into fundamental theories in operations management. One important result in this research is that there is strong evidence that the ISO 9001 registration efforts enhance organizational competitiveness. This is a significant contribution due to the fact that there is substantial disagreement in the literature over the relationship between the two constructs (Brecka, 1994; Stratton, 1994; Anderson et al., 1999).

The results of this study agree with Anderson et al. (1999) who indicated that customer and regulatory compliance are not the only reasons for the widespread adoption of ISO 9000 in North

American manufacturing companies. They found that firms pursue ISO 9001 as a means of achieving quality improvement and global competitiveness.

The results of the analysis support the belief that TQM contributes to competitiveness (Reimann and Hertz, 1994; Knod, Jr. and Schonberger, 2001). ISO 9000 registration efforts and TQM practices have a significant, positive relationship when the model is tested based on all 295 responses. The overall result supports the claim that ISO 9000 might be a good first step to total quality and is a meaningful component of TQM. Conflicting conclusions, however, can be drawn, depending on the firm size. Small firms seem to side with critics who claim that ISO 9000 has little relation to TQM. As a matter of fact, our findings support the claim of McAdam and McKeown (1999), which states, although many small businesses in Northern Ireland are benefiting from ISO 9000, the majority of firms are not progressing towards TQM. On the other hand, larger firms tend to side with proponents who see ISO 9001 as a starting point for TQM and as an ongoing integral part of TQM.

Given these findings, it becomes clear why the term ISO 9001 has become more widely used than TQM in discussions of quality improvement and global competitiveness. Since ISO 9001 is much smaller in scale than TQM, it is more manageable and achievable, and yet, ISO 9001 registration efforts can give similar benefits to the organization even though they may not be in the same scale. Furthermore, companies that are ISO 9001 certified receive acknowledgement and recognition from a third party as well as their customers.

This study was developed from a theoretical foundation and provides a deeper insight into fundamental theories in operations management. One important result in this research is that there is strong evidence that the ISO 9001 registration efforts enhance organizational performances. The results of the analysis also support the belief that ISO 9001 contributes to competitiveness, which in turn helps to gain greater profitability.

CONCLUSIONS

In this study, SEM was applied to help explain and predict the relationships between and among ISO 9001 registration efforts, TQM practices, firm performance. A model framework was formulated based on the existing literature. Three hypotheses were proposed and tested based on the empirical data collected from ISO 9001 registered companies in the United States, and the results showed consistent support for 2 of the 3 hypotheses.

This study has contributed to the existing literature on quality management in several ways. First, it identified significant research issues and addressed previously unanswered questions. Second, a model was formulated and introduced based on the existing literature to provide a clearer understanding of the relationships among existing constructs. Third, this empirical study supported or refuted existing beliefs and propositions, as well as fostered development of new theories and concepts in quality and operations management.

REFERENCES

References are available from S. Bruce Han upon request.

Table 1: Reliability and convergent validity of constructs.

Construct and Indicators	Standardized Loading	Reliability
ISO 9000		.85
Quality System	.69	
Document and Data Control	.69	
Process Control	.63	
Corrective and Preventive Action	.69	
Quality Records	.71	
Internal Quality Auditing	.75	
Training	.62	
Range of t-values	9.23 – 10.08	
TQM Practices		.82
Supplier Relationship	.59	
Customer Involvement	.61	
Training	.86	
Top Management Commitment	.76	
Product Design	.76	
Range of t-values	9.72 – 14.51	
Firm Performance		.88
Cost	.83	
Quality	.69	
Delivery	.84	
Flexibility	.84	
Range of t-values	12.77 – 16.15	

Note: All t tests were significant at $p<.001$

Table 2: Correlation matrix – constructs

	ISO 9001	TQM	Firm Performance
ISO 9001	1.000		
TQM	0.190	1.000	
Firm Performance	0.407	0.528	1.000

Note: All were significant at $p<.01$

Table 3: Multiple group analyses.

Hypotheses	ALL		FIRM SIZE Small		Large	
H1:	.19*	yes	-.02	no	.27**	yes
H2:	.32**	yes	.23*	yes	.35**	yes
H3:	.47**	yes	.54**	yes	.44**	yes
Chi-Sq.:	176.8		127.0		142.5	
df:	101		101		101	
Ratio:	1.75		1.26		1.41	
p-value:	.000		.041		.004	
CFI:	.994		.992		.996	
TLI:	.993		.989		.995	
RMSEA:	.051		.060		.043	
n:	295		72		223	

Note: *: p<.01; **: p<.001
n = Sample Size

ANAYLYZING THE IMPACT OF OPERATIONS MANAGEMENT ON FINANCIAL PERFORMANCE – A CRITICAL REVIEW

Beate Klingenberg
Marist College
School of Management
Poughkeepsie, NY 12603
+1-845-575-3000 x6071
Beate.Klingenberg@marist.edu
Rachel Timberlake
Ultra Seal Corporation
845-392-1909
Rachel.timberlake@gmail.com
Tom G. Geurts
New York University
Schack Institute of Real Estate
11 West 42nd Street, Room 509-A
New York, NY 10036-8083
Tel: (212) 992-3241
Tom.Geurts@nyu.edu

ABSTRACT

The purpose of operations is the management of an organization's processes to increase productivity by using methods such as Just in Time (JIT)/Lean Manufacturing, Total Quality Management (TQM), Environmental Management Systems (EMS), or Six Sigma. Following the implementation of such methods, managers want to determine their impact on performance. Past studies have analyzed financial ratios to prove the usefulness of the operational methods. This paper critically reviews the appropriateness of profitability ratios to determine financial performance given a specific operations strategy. Focusing on JIT/Lean Manufacturing, the correlation between profitability ratios and ratios that reflect inventory management policies is analyzed.
Keywords: Financial Performance, Just-in-Time Manufacturing, Lean Manufacturing, Profitability, Inventory Management Ratios

INTRODUCTION

The field of operations management continuously sees innovations in the management of the production process. Long established ones focus on continuous improvement of quality in general – such as Lean Manufacturing or Just-In-Time (JIT) and Total Quality Management (TQM) while more recently, environmental quality and sustainability are in the focus through Environmental

Management Systems (EMS), "greening" of operations, or Lean Green Six Sigma. The academic literature provides various studies that analyze the impact of these methods on the performance of firms that have implemented them, focusing on financial performance through various financial ratios, such as ROA, ROE, profit margin, etc. [1-13], or building more sophisticated models that also include metrics of customer satisfaction and competitiveness [14]. This paper takes a renewed look at which financial ratios are best suited to analyze a change in the financial performance given one the innovative operations management strategies discussed above. The focus of this paper is on JIT. The following section discusses financial ratios that appear appropriate to study JIT impact and their relationship to firm profitability, leading to hypotheses that are tested in the third section, followed by a discussion and further research outlook.

MEASURING THE IMPACT OF JUST IN TIME/LEAN MANUFACTURING[1]

JIT is a complex philosophy for process management, with the primary goal of elimination of waste through the simplification of the manufacturing process. Waste, as defined in JIT, is anything more than the absolute minimum necessary in order to complete a product or task.
JIT systems have the ability to produce large quantities of fairly customized products just as they are demanded by the end user on resources similar to those employed for mass production of limited variety output utilizing traditional manufacturing techniques. JIT systems require less work in process inventory to achieve high levels of efficiency, lowering per unit costs. Additionally, producing at the customer's rate of consumption produces less finished goods inventory, incurring less cost to carry.
Since JIT affects the inventory on the balance sheet, the financial ratios that include the inventory in its calculation should be indicators of successful implementation.
The following widely used financial rations employ inventory:

Inventory to Current Assets= (Inventory/Current Assets) * 100

This ratio gives the inventory as a percentage of the current assets and as such measures the percentage of current assets tied up in inventory. Since inventories are not self liquidating a high Inventory to Current Assets can indicate obsolete or defect inventory. Clearly, after introducing JIT this number should decrease significantly, theoretically to zero. Since there are costs associated with carrying inventory this would indicate that these costs are reduced.

Inventory Turnover= (Cost of Goods Sold/Average Inventory) * 100

This ratio evaluates the operating efficiency with regards to the usage of inventory in the production process. Sometimes net sales are used in the numerator. Regardless, since JIT theoretically reduces the denominator to zero, this ratio is no longer measurable. Hence, this phenomenon will indicate the correct implementation of JIT.
Finally, one can evaluate the difference between the Current and Quick Ratio:

Current Ratio= Current Assets/ Current Liabilities
Quick Ratio= Current Assets-Inventory/Current Liabilities

[1] For readability of the paper, JIT/Lean manufacturing is only referred to as JIT.

JIT implementation reduces inventory, and hence the Current Ratio should be equal to the Quick Ratio.
Clearly, financial ratios that measure inventory directly should reflect the implementation of JIT. The question remains, however, how JIT affects overall firm performance, in particular profitability. Profitability ratios such as Return on Assets (ROA = net sales/total assets) and Return on Equity (ROE = net sales/common equity), however, also include effects of financial leverage. If anything, then a metric must be used that is not impacted by financial leverage, although it will still include all other firm activities that impact profitability. Such a metric is Basic Earning Power (BEP = EBIT/Total assets; with EBIT = Earnings before Interest and Taxes).
Rather than measuring the changes of inventory and profitability ratios before and after JIT implementation as most prior studies have done, it is suggested to first test the relationship between these ratios to determine the validity of their use in such an analysis. This leads to the following hypotheses:

H1: A correlation exists between BEP and ratios that reflect inventory
H2: No correlation exists between ROE and ratios that reflect inventory
H3: No correlation exists between ROA and ratios that reflect inventory.

These hypotheses assume that the reduced cost through JIT should at least be reflected in a profitability ratio that is not influenced by the leverage effect – thus BEP - whereas the ratios ROA and ROE do not clearly reflect this impact as they also include the effects of financial leverage. For hypothesis H1, a one-tailed test is used; whereas H2 and H3 are using a two-tailed test.

PRELIMINARY RESULTS

The automotive industry is chosen to test the hypotheses formulated in the previous section. This industry is considered the "poster child" for operations management innovations such as JIT or TQM: in fact JIT/Lean manufacturing was to a large extent developed by Toyota through the Toyota Product System. It is assumed that most firms in this industry use JIT at some level, and as the focus of this study is determining the relationship between inventory and profitability ratios, it is irrelevant to which extend a firm uses JIT practices.
Toyota, Ford and General Motors (GM) are chosen as industry representatives. Financial Data from Income Statements and Balance Sheets is taken from the Thomson Analytics Database. The following financial ratios are either directly taken from the database, or calculated based on reporting: inventory/current assets, current and quick ratio as well as the difference (current – quick), inventory turnover, ROA, ROE and BEP.

Hypothesis Testing

The hypotheses are further detailed as follows:

H1.1: The BEP is positively correlated to Inventory Turnover
H1.2: The BEP is negatively correlated to Inventory/current assets
H1.3: The BEP is negatively correlated to Current-Quick

H2.1: The ROE is not correlated to Inventory Turnover
H2.2: The ROE is not correlated to Inventory/ current assets
H2.3: The ROE is not correlated to Current-Quick

H3.1: The ROA is not correlated to Inventory Turnover
H3.2: The ROA is not correlated to Inventory/ current assets
H3.3: The ROA is not correlated to Current-Quick

The rationale for BEP is as follows: high inventory turnover should result in increased sales and can lead to increased profitability if costs are kept equal; low percentage of inventory in current assets should improve profitability through cost decreases; low inventory levels as represented by a small difference between current and quick ratio should improve profitability through cost savings.
Table 4 summarizes the result.

Table 4: Hypothesis testing for Toyota, Ford and GM. Numbers in italics indicate significant correlations for p>0.5.

	Toyota	Ford	GM
H1.1 Positive Correlation BEP vs. Inventory Turnover	*0.437*	0.138	*0.647*
H1.2 Negative Correlation BEP vs. Inventory/Current Assets	-0.124	0.229	*-0.496*
H1.3 Negative Correlation BEP vs. Current-Quick	0.200	-0.146	*-0.295*
H2.1 No Correlation ROE vs. Inventory Turnover	*0.368*	-0.207	-0.141
H2.2 No Correlation ROE vs. Inventory/Current Assets	-0.074	-0.170	0.154
H2.3 No Correlation ROE vs. Current-Quick	0.229	0.221	0.178
H3.1 No Correlation ROA vs. Inventory Turnover	*0.412*	*0.273*	*0.495*
H3.2 No Correlation ROA vs. Inventory/Current Assets	-0.114	*0.294*	*-0.404*
H3.3 No Correlation ROA vs. Current-Quick	0.205	-0.119	*-0.325*

Significant correlations supporting H1.1 (The BEP is positively correlated to Inventory Turnover) are found for both Toyota and GM, however not for Ford, although the sign of the correlation is correct. H1.2 (The BEP is negatively correlated to Inventory/current assets) is supported by GM, with a correct sign of the correlation coefficient for Toyota, but not for Ford. H1.3 (The BEP is negatively correlated to Current-Quick) is by the GM data, however, the sign of the correlation for Toyota is not as expected. Based on these preliminary correlations, it cannot be concluded with absolute certainty, but with high probability, that BEP reflects inventory policies.
H2.1 (The ROE is not correlated to Inventory Turnover), H2.2 (The ROE is not correlated to Inventory/ current assets) and H2.3 (The ROE is not correlated to Current-Quick) is supported by the data, except that there is a correlation between ROE and Inventory Turnover for Toyota. These results support that ROE is probably not an appropriate ratio to measure the impact of inventory management.
H3.1 (The ROA is not correlated to Inventory Turnover) is rejected for all firms, and there is a correlation between inventory/current assets and ROA (H3.2) for Ford and GM, as well as one between current-quick and ROA for GM.

It appears therefore that there is some merit in using ROA to analyze the impact of inventory management techniques, such as JIT. The correlation is probably based on the fact that ROA contains total assets in the denominator, and hence is directly impacted by decreasing inventory levels.

DISCUSSION AND OUTLOOK

The analysis attempts to determine how the impact of JIT on financial firm performance, in particular on profitability can be measured. The academic literature offers studies that compare for example ROA and ROE over the time of JIT implementation to measure the success of this activity. However, ROA and ROE also reflect other managerial decisions (investment activities, other operational activities, financial leverage) besides inventory policies, which can “dilute” or even reverse the expected positive impact of JIT. Before using ROA and ROE, it is useful to determine if both ratios indeed correlate to inventory measures as reflected in inventory management ratios. Additionally, an alternate ratio, BEP, is suggested as at least the leverage effect can be excluded.

The preliminary results indicate that ROE is indeed not a suitable metric to determine the impact of inventory management methods on firm profitability. As for ROA and BEP, the results thus far are inconclusive, although encouraging enough to continue research efforts.

Future work will broaden the analysis to other firms in the industry as well as to other industries. Although the automotive industry represents a “poster child” for JIT implementation, large corporations such as Ford and GM do not only assemble and sell cars, but also have significant income streams through their financing divisions. The available financial data, onto which this analysis is based, however, does not allow separating this income, its impact on ROA, ROE and BEP, which could potentially skew the results. Other manufacturing industries that do not diversify into non-manufacturing activities will therefore be included.

Furthermore, a more detailed analysis of the reported financial data will be added. Although all three firms chosen in this analysis are required to report according to GAAP principles, or to reconcile their filings according to these principles (in the case of Toyota), differences can still occur and may result in different reporting of inventory or other items on income statements and balance sheets [15]. Such an analysis might explain why the signs of correlations are - in a few incidents - not as expected.

To conclude, even in this early stage, the results are adding interesting points of discussion to the academic literature and encourage a critical discussion of the applicability of financial ratios for the measurement of operational innovations.

REFERENCES

[1] Ahmad, A., Mehra, S., Pletcher M., “The perceived impact of JIT implementation on firm’s financial/growth performance”, 2004, *Journal of Manufacturing Technology and Management*, 15(2), pages. 118-130.

[2] Fullerton, R.R., McWatters, C.S., Fawson C. “An examination of the relationships between JIT and financial performance”, 2003, *Journal of Operations Management*, 21(4), pages 383-404.
[3] Kaynak, H., “The relationship between just-in-time purchasing techniques and firm performance”, 2002, *IEEE Transactions on Engineering Management*, 49(3), p. 205.
[4] Boyd, D.T., Kronk, L., Skinner, R., “The effects of just-in-time systems on financial accounting metrics”, 2002, *Industrial Management + Data Systems*, 102(3/4), pages 153-165.
[5] Benner, M.J., Velosos, F.M., “ISO 9000 practices and financial performance: A technology coherence perspective, 2008, *Journal of Operations Management*, 26(5), page 611.
[6] Wayhan, V.B., Balderson, E.L, “TQM and Financial Performance: A Research Standard”, 2007, *Total Quality Management & Business Excellence*, 18(4), page 393.
[7] Kaynak, H., “Measuring organizational effectiveness and business performance in firms implementing total quality management”, 2006, *International Journal of Manufacturing Technology and Management*, 8(4), page 355.
[8] Kaynak, H., “The relationship between total quality management practices and their effects on firm performance”, 2003, *Journal of Operations Management*, 21(4), pages 405-436.
[9] Eriksson, H., Hansson, J., “The impact of TQM on financial performance”, 2003, *Measuring Business Excellence*, 7(1), Pages 36-51.
[10] Borri, F. and Boccaletti G. “From total quality management to total quality environmental management”, *The TQM Magazine*, 1995, 7(5), pages 38-42.
[11] Stead, W.E. and Stead J.G., *Management for small planet. Strategic decision making and the environment*. Newberry Park, California, Sage Publications, 2003.
[12] Ward, M. “Life cycle: The preferred environmental strategy”, 1994, *Chemical Week*, 154(16), page 23.
[13] Reference removed for the confidentiality of the review process.
[14] Han S.B, Chang, S.K., Ebrahimpour, M. "The impact of ISO9000 on TQM and business performance." *The Journal of Business and Economic Study*, 2007, 13 (2), pages 1-25.
[15] Kennedy, B.B., Kennedy, K.,. Olinsky, A. "A study of the differences in financial statements prepared using U.S. GAAP and international financial reporting standards for three major European auto manufacturers". Proceedings of North Eastern Decision Sciences Annual Meeting, Uncasville, CT, 2009.

SPREADSHEET MODELING WITH VBA FOR THE SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

Jae-Dong Hong, Industrial Engineering Technology, South Carolina State University, Orangeburg, South Carolina 29117, (803) 536-8861, jhong@scsu.edu
Yuanchang Xie, Civil Engineering Technology, South Carolina State University, Orangeburg, South Carolina 29117, (803) 536-8321, yxie@scsu.edu
KiYoung Jeong, Engineering Management & Systems Engineering, University of Houston at Clear Lake, Houston, TX 77058, (281)283-3862, jeongk@uhcl.edu

ABSTRACT

Traffic assignment involves assigning traffic to a transportation network typically consisting of highways, arterials, or transit routes. To obtain an optimal traffic assignment scheme, a mathematical program (MP) should be formulated and solved. The process of formulating an MP requires identifying all possible paths connecting each origin to destination pair through the network. Consequently, formulating and solving such an MP is a very difficult task due to the large number of decision variables and constraints. In this paper, we suggest a new and efficient way of formulating an MP and develop an Excel model with VBA. The developed Excel model can be easily expanded to a large-scale transportation network.

Keywords: Traffic Assignment, Transportation Network, Mathematical Program, VBA

INTRODUCTION

Traffic Assignment Background

Traditional transportation planning consists of the following four steps: trip generation, trip distribution, mode split, and trip assignment (Horvath [1]). The entire planning area is divided into a number of Traffic Analysis Zones (TAZs). The results from the first three steps are Origin-Destination (O-D) matrices showing the travel demand among those TAZs in the future. Each O-D matrix represents one type of trips, for instance, passenger-car trips. The trip assignment step is to assign the projected travel demand onto the transportation network so that potential problems such as congestion can be identified. In other words, trip assignment is to predict the number of travelers using various routes and, hence, the traffic volumes on the links of a transportation network. When vehicular trips rather than person trips are estimated, we call it a *traffic assignment problem* (Banks [2, p. 368]).

The fundamental aim of a traffic assignment process is to predict or reproduce the pattern of vehicular movements that would be observed when the travel demand represented by the O-D matrix to be assigned is satisfied. The major aims of traffic assignment procedures are shown in Matthew [3, p. 1]. Typical traffic assignment methods include all-or-nothing assignments, user equilibrium assignments, and system optimum assignments. An all-or-nothing assignment is often referred to as the minimum path or shortest path algorithm. The shortest path, or tree, represents the shortest time path between the centroids of two TAZs and is assigned all of the traffic volume between the two TAZs in question. This method is unrealistic as only one path between every O-D pair is utilized, even if there may be another path with the same travel time or cost. In addition, this method ignores the fact that link travel time is a function of link traffic volume. As link traffic volume increase, link travel time will increase accordingly. The initial shortest paths will no longer be the shortest ones and travelers will switch to other less congested routes. However, this important effect is not taken into account in the all-or-nothing assignment. Thus this assignment method often generates unrealistic flow patterns when link capacity constraints and congestion effects exist.

The user equilibrium (UE) assignment is based on Wardrop's [4] first principle, which states that no traveler can improve his travel time by unilaterally changing routes. This principle is also called Wardrop's selfish equilibrium. In contrast to UE, the system optimum (SO) assignment is based on Wardrop's [4] second principle, which states that drivers cooperate with one another in order to minimize total system travel time. This assignment can be thought of as a model in which congestion is minimized when drivers are told which routes to use. The SO assignment can be useful to traffic engineers trying to manage the traffic to minimize the system travel costs and therefore achieve an optimum social equilibrium. In the present paper, we will consider the SO assignment problems.

Mathematical Programming Approach

Let G= (N, A) denote the transportation network consisting of N nodes and A directed links. And let

f_k^{rs} = Flow on path k connecting O-D pair r-s.

k = Index for the specific path.

q_{rs} = Traffic demand between the origin 'r' and destination 's.'

$t_a(x_a)$ = Travel time on link a for a given link flow, x_a.

x_a = Flow on link a, $a \in A$.

$\delta_{a,k}^{rs}$ = An indicator variable: $\delta_{a,k}^{rs}$ = 1, if link a is a part of path k connecting O-D pair r-s; otherwise $\delta_{a,k}^{rs}$ =0.

To find the link flows that satisfy the SO criterion when all the O-D entries have been appropriately assigned, the following integer nonlinear program (INLP) can be formulated (Sheffi [5, pp. 69-70])

$$Minimize \;\; Z = \sum_a x_a t_a(x_a), \tag{1}$$

subject to

$$\sum_k f_k^{rs} = q_{rs}, \; \forall r \; and \; s, \tag{2}$$

$$x_a = \sum_r \sum_s \sum_k \delta_{a,k}^{rs} f_k^{rs}, \forall a, \tag{3}$$

$$f_k^{rs} \geq 0, \;\; \forall r \; and \; s, \tag{4}$$

$$x_a \geq 0, \;\; a \in A, \tag{5}$$

where Equation (2) is known as *path-link incidence relationships* and Equation (3) is called *flow conservation constraints*. Sheffi [5] pointed out that the SO assignment can result only from the cooperative decisions by all motorists to minimize the total system travel time rather than their own. Thus the SO flow pattern is not commonly observed in the real world. The significance of the SO formulation is that its solution may serve as a benchmark based on which different flow patterns can be compared.

A NEW METHOD OF FORMULATION

Phase I

In Phase I, we use x_{ij} to denote $x_a \geq 0,\ a = (i, j)\ , \forall (i, j) \in N$. Let K^{rs} denote the number of paths connecting O-D pair *r-s*. Define an indicator variable, $\zeta_k^{rs} = 1$, if K^{rs} =1; otherwise, $\zeta_k^{rs} = 0$. The decision variable in a new formulated mathematical program is the flow from node *i* to node *j*, x_{ij} ($i \neq j$), rather than f_k^{rs}. Note that the SO problem can be transformed into a network flow problem subject to the net-flow to node *i* equaling the net-traffic demand between node *i* and node *i + u*, *u = 1, 2, …n-i*. In addition, if only one path between node *i* and node *j* exists in the SO traffic assignment problem, then x_{ij} should be greater than or equal to, q_{ij}, the traffic demand between node *i* and node *j*, in order to satisfy the required traffic demand between node *i* and node *j*. Now, the SO problem would be reformulated as follows:

$$Minimize\ \ Z = \sum_i \sum_j x_{ij} t_{ij}(x_{ij}), \tag{6}$$

subject to

$$\sum_u (x_{iu} - x_{ui}) = \sum_u (q_{iu} - q_{ui}),\ \forall i, u \in N, \tag{7}$$

$$x_{ij} \geq q_{ij} \zeta_k^{ij}, \forall k, i\ and\ j, \tag{8}$$

$$x_{ij} \geq 0,\ \forall (i, j) \in N. \tag{9}$$

The INLP problem in Equations (6)-(9) yields the same optimal objective solutions, Z^* and x_{ij}^*, as the problem given in (1)-(5). The left-hand side (LHS) of Equation (7) denotes the net-flow to node *i* resulting from subtracting the incoming flow from the outgoing flow, whereas the right-hand side (RHS) is the net-traffic demand for node *i*. Thus Equation (7) represents Balance-of-Flow Rules. Note that the number of decision variables is the same as the number of links, not the number of paths; so the number of decision variables and the corresponding constraints would be significantly reduced.

Phase II

Once x_{ij}^* is obtained from solving the MP given in (6)-(9), f_k^{rs} can be obtained by simultaneously solving Equation (2) and

$$\sum_r \sum_s \sum_k \delta_{a,k}^{rs} f_k^{rs} = x_{ij}^*,\ a = (i, j), \forall a. \tag{10}$$

To reduce the computational effort, we could delete the f_k^{rs} for $q_{rs} = 0,$ from Equations (2) and (10) by setting $f_k^{rs} = 0, \forall k.$ In addition, we could also delete the f_k^{rs} for $\zeta_k^{ij} = 1$ for a given *i* and *j*, since we can set $f_k^{ij} = x_{ij}^*.$ In the following, we develop instead an Excel spreadsheet model with VBA code to solve this two-phase model.

DEVELOPMENT OF A SPREADSHEET MODEL WITH VBA CODE

Our spreadsheet model is not fixed in size, but determined by the total number of nodes. Since the number of decision variable depends upon the number of links, without VBA coding it would be very difficult to find the optimal solution using the spreadsheet model alone for a general SO traffic assignment problem. We use an example network with five nodes as shown in Figure 1. The corresponding O-D flow matrix is given in Table 1. We use the link cost suggested by the US Bureau of Public Road (BPR), which is given by

$$t_a(x_a) = t_a^o\left[1 + \alpha(x_a / C_a)^\beta\right], \tag{11}$$

where t_a^o is the travel time per unit distance under free flow conditions for link a, C_a is the practical capacity of link a, and α and β are parameters to be calibrated. In this paper, we assume α=1 and β=1. Based on data in Tables 1 and 2, we use a spreadsheet procedure, solve the problem using the Premium Solver function, and present the solution for x_{ij}^* and Z^* in Figure 2. The total travel time, Z^*, is the optimal value of the objective function given by Equation (1), but the optimal flow on path k connecting O-D pair r-s, f_k^{rs}, should be obtained from the optimal link flow x_{ij}^* in Phase II. Figure 3 shows the optimal flow on each path.

SUMMARY, CONCLUSIONS, AND FUTURE RESEARCH

In this paper, we consider a system optimal traffic assignment problem. The objective is to minimize the total system travel time, subject to satisfying the traffic demand between each origin to destination pair. The problem has traditionally been formulated as a nonlinear integer programming problem with a large number of decision variables and constraints, which makes it very difficult to solve. We propose a simplified two-phase method of formulating the problem and develop a Microsoft Excel model with VBA code to solve this problem with much less efforts. As illustrated in the previous sections, the traditional method of formulating it as an MP problem requires twenty three decision variables and sixteen constraints, whereas our proposed two-phase method requires only eight decision variables and nine constraints in Phase I and requires simultaneously solving nine equations with thirteen variables in Phase II. The resulting problem is solved by the premium Solver function in Excel.

Figure 1. Eight-Link Network.

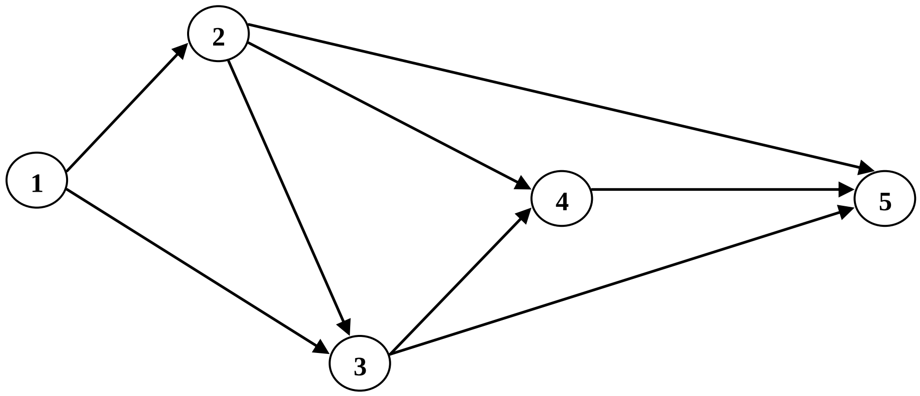


Figure 2. Optimal Solution to Link Flow

	A	B	C	D	E	F	G	H	I	J	K	L	M
1				**Free-Flow Time Matrix**									
2					Destination								
3		Origin	1	2	3	4	5						
4		1		10	15								
5		2			3	5	9						
6		3				4	7						
7		4					10						
8		5											
9													
10													
11				**Capacity Matrix**				Traffic Assignment					
12					Destination								
13		Origin	1	2	3	4	5						
14		1		300	500								
15		2			150	200	150						
16		3				200	200						
17		4					150						
18		5											
19													
20													
21				**Flow Matrix**									
22					Destination								
23		Origin	1	2	3	4	5	Total					
24		1		100	100	100	200						
25		2			50	50							
26		3				100							
27		4					50						
28		5											
29		Total											
30													
31	No.	Origin	Destination	Link Flow	Time	Capacity	Zeta	lb Flow	Nodes	NetFlow	=	NetDemand	
32	1	1	2	256	18.5333	300	1	100	1	500	=	500	
33	2	1	3	244	22.32	500	0	0	2	0	=	0	
34	3	2	3	50	4	150	1	50	3	50	=	50	
35	4	2	4	129	8.225	200	0	0	4	200	=	200	
36	5	2	5	77	13.62	150	0	0	5	250	=	250	
37	6	3	4	121	6.42	200	1	100					
38	7	3	5	123	11.305	200	0	0					
39	8	4	5	50	13.3333	150	1	50					
40													
41				Total Travel Time	15334.4								

Figure 3. Spreadsheet Modeling for Finding the Optimal Path Flows

Path No.						Path Flow	Path-Link Flow Matrix 1-2	1-3	2-3	2-4	2-5	3-4	3-5	4-5		Path-Demand Matrix 1-2	1-3	1-4	1-5	2-3	2-4	3-4	4-5
1	1	2				100	1									1							
2	1	3				100		1									1						
3	2	3				50			1											1			
4	2	4				50				1											1		
5	2	5				0					1												
6	3	4				100						1										1	
7	3	5				0							1										
8	4	5				50								1									1
9	1	2	3			0	1		1								1						
10	1	2	4			79	1			1								1					
11	1	2	5			77	1				1								1				
12	1	3	4			21		1				1						1					
13	1	3	5			123		1					1						1				
14	2	3	4			0			1			1									1		
15	2	3	5			0			1				1										
16	2	4	5			0				1				1									
17	3	4	5			0						1		1									
18	1	2	3	4		0	1		1			1						1					
19	1	2	3	5		0	1		1				1						1				
20	1	2	4	5		0	1			1				1					1				
21	1	3	4	5		0		1				1		1					1				
22	2	3	4	5		0			1			1		1									
23	1	2	3	4	5	0	1		1			1		1					1				
						Flow	256	244	50	129	77	121	123	50	Flow	100	100	100	200	50	50	100	50
							=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=
						Required Flow	256	244	50	129	77	121	123	50	Required Demand	100	100	100	200	50	50	100	50

Path No	Path Flow	Reduced Path-Link Flow Matrix 1-2	1-3	2-4	2-5	3-5	Reduced Path-Demand Matrix 1-3	1-4	1-5	2-4
2	100		1				1			
4	50			1						1
9	0	1					1			
10	79	1		1				1		
11	77	1			1				1	
12	21		1					1		
13	123		1			1			1	
14	0									1
18	0	1						1		
19	0	1				1			1	
20	0	1		1					1	
21	0		1						1	
23	0								1	
	Flow	156	244	129	77	123	100	100	200	50
		=	=	=	=	=	=	=	=	=
	Required Flow	156	244	129	77	123	100	100	200	50

Table 1. O-D Flow Matrix

	To				
From	1	2	3	4	5
1	0	100	100	100	200
2	0	0	50	50	0
3	0	0	0	100	0
4	0	0	0	0	50
5	0	0	0	0	0

Table 2. Link Characteristics (t_o^a, C_a)

	To				
From	1	2	3	4	5
1		(10,300)	(15, 500)		
2			(3, 150)	(5,200)	(9, 150)
3				(4, 200)	(7, 200)
4					(10, 150)
5					

REFERENCES

References available upon request from the authors

Behavior of a Remanufacturing System in Presence of Unreliable Suppliers for New Product

Suman Niranjan* and Amit Vasavada**
*School of Management, State University of New York - Binghamton, Vestal, NY 13902
**System Science and Industrial Engineering, State University of New York – Binghamton Vestal, NY 13902

1. Introduction

Products or components are remanufactured when they are returned before their end of life. Remanufacturing falls under the broad area of closed loop supply chain and reverse logistics. Closed loop supply chains (CLSC) have been receiving a lot of attention in the industry and academia over last two decades. CLSC focuses on several aspects such as remanufacturing the product before its end of life, reusing certain modules of the returned product, and recycling products after their end of life. In all cases the products are either returned by the customers to the manufacturer or the manufacturer collects them through different channels. Remanufacturing has a huge economic potential, this sector accounts for $53 billion in terms of sales and employment [6]. There are several return types as mentioned earlier, but in this paper we focus on the products that are returned before their end of life cycle. For instance, if the products are returned early in their life cycle, the returned product can be refurbished and in turn used to fill the warranty demands [3].

Nike is one of the world's largest footwear manufacturers recycles about 125,000 pairs of shoes in the continent of U.S; it has recycled about 15 million pairs of shoes worldwide since 1993 [6]. Nike recycles and reuses 100% of the materials from the consumers returned shoes and uses it for making sport surfaces. Xerox Corporation is one of the most widely reported remanufacturing systems in the world [4]. Xerox manufactures products related to printing, publishing, storing and sharing documents. Xerox has recovered products since 1960s, but it was in late 1980s and early 1990s when the remanufacturing operations were used for profitability. The products which return to the company at end of life or end of lease are inspected, and good components from the product are either used in the new products or used as spare components if the product is no longer used in new product manufacturing. IBM's most attractive business has been end of lease equipment, leases at IBM account for 35% of IBM's hardware output [2]. IBM extracts the spare parts from its products that are not working and uses them to repair defective parts in similar or same products. 80% of the mobile phone users in United States upgrade their perfectly working mobile phone every year; this results in the product being returned before its end of life [3], the perfectly working components from the mobile phone can be used in new products or as spare parts, this is true for several other consumer electronic items.

In this situation an important research question arises, what should be the right mix of new and remanufactured components used in the manufacturing of a new product, and why should we care about the mix? If higher volumes of used but perfectly working components are placed in the manufacturing of the new product, it can reduce the life of the new product putting the manufacturer's warranty at stake. Using lower volumes of remanufactured product will result in accumulation of these components. There is certain amount of cost that are associated with the remanufactured products, like the value of the product (purchase price), inspection cost, remanufacturing/refurbishing cost etc. This suggests that the longer you hold these products in the inventory the higher would be their holding costs.

In this paper we study a two-echelon remanufacturing system, which focuses on the research question that we discussed earlier. Moreover we study this problem in the presence of unreliable suppliers. Specifically we look at a component that is used in the assembly of new product. This component can either be a remanufactured component or a new component. The new component is procured from suppliers whose supply reliability is not 100%. The recovered and remanufactured product is assumed to be perfect substitute for the new product. The end product demand is stochastic in nature, each node has stochastic capacity utilization, there is a fixed manufacturing/assembly lead time between the nodes, and the arrivals of the returned product are random in nature. An installation type periodic base-stock policy is used for the two-echelon system. To model the return flows, we let returns depend on the end product demand, a fraction of the end product demand is assumed to return. We analyze the performance of the system using a simulation based optimization approach in OptQuest (an optimization tool in ARENA, simulation software). Due to a number of realistic assumptions developing an analytically tractable model would be very difficult, so we resort to simulation optimization. The computational study was conducted for several interesting cases using a computational study. Specifically we try to understand the impact of the following i) best possible scenario under a given situation, based on lowest optimal supply chain cost ii) behavior of the system under random supply reliability for new components, and iii) optimal target stock level for new and remanufactured components.

The rest of the paper is organized as follows. In section 2 we introduce our model and section 3 we describe the cases followed by inferences from the numerical results, and we conclude in section 4 with future scope and directions.

2. Model

Figure 1 show the schematic of two-echelon remanufacturing system, where node 2 is a remanufacturing unit, which procures the returns from the customers and other sources, the returns for the product is assumed to be a function of the final product demand. A yield is associated with node 2, we assumed that not all products that are returned are successfully recovered, due to two reasons i) components from the returned product not usable, ii) the cost of remanufacturing outweighs the cost of using a new component. Node 1 is a warehouse which holds the remanufactured components and new components, the new components are procured from suppliers whose supply reliability varies. Node 0 manufactures a final product, which uses either the remanufactured or new components procured from node 1. A fixed lead time is associated between the echelons, which corresponds to the ordering lead time and manufacturing lead time. Uncertainty is involved in demand, capacity and yield. Uncertainty is also involved in the cost of the new component supplied by the suppliers.

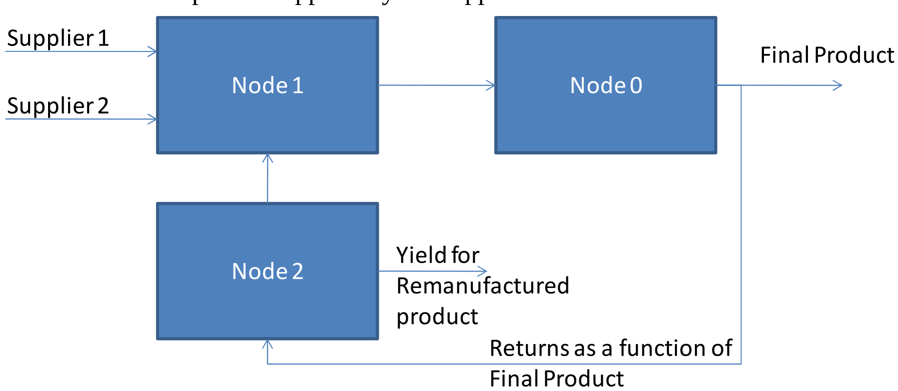


Figure 1: Two-echelon Remanufacturing System

To describe the operations through the rest of the paper the following notations are used:-

ξ_n^j : Product *j* demand in period *n*

η_n^i : Realized capacity at stage *i* in period *n*

s^i : Base-stock level for item *i*

c^i : Cost per unit of item *i*

Y_n^i : Outstanding orders of item *i* in period *n* that have not been delivered

i_n^i : On-hand inventory level of item *i* in period *n* before the demand is realized

IL_n^i : Inventory level of the item *i* in period *n* before the demand is realized

DS_n^i : Downstream shortage at node *i* in period *n*

ERU : Excess recycled units

ENU : Excess new units

IRU : Initial recycled units

INU : Initial New units

α^i : Required type-I service level at node *i*

ρ : Fraction of demand coming from the recycled products

ε_n : Yield of returned goods at Node 2

3. Two-echelon Remanufacturing System Assumptions and Operations

The two-echelon remanufacturing system considered here operates under a periodic-review installation base-stock policy, with inventory located after the processing at the node. Under this policy, at the start of each period if the inventory position (*on-hand inventory + orders – backorders*) falls below the base-stock level an order is placed to bring the inventory position back to the base-stock level. Unsatisfied orders are backlogged. In order to ensure system stability, we assume that that the average demand is less than the average capacity at a node $E\left[\xi_n^i\right] < E\left[\eta_n^i\right]$, except for node 2 whose demand is only a fraction of the final product demand, the fraction is equal to ρ, so the stability assumption for node 2 would be $\rho * E\left[\xi_n^i\right] < E\left[\eta_n^i\right]$. In order to ensure reliability of the final product, we consider only a fraction (ρ) of the components used in the manufacturing of the final product would come from the remanufactured product and the rest would come from the new components. At the beginning of each period the following sequence of activities occur: i) the outstanding orders are updated, (i.e. items that have not been delivered in the previous period are accounted for), ii) the on-hand inventory is updated, (i.e. the physical inventory), iii) demand is realized iv) capacity is realized. The objective function in (1) indirectly penalizes holding inventory at each location, since higher s^i corresponds to more inventory of item *i* [1]. Apart from the cost of the product The constraints are formulated on the basis of a type-I service level. The constraints ensure that sufficient inventory is held to meet demands with a high level of certainty.

Since node 1 corresponds to warehouse which stocks both remanufactured In equation 1, *1n* and *1s* correspond to node 1. These notations are used for analytical reasons, where *1n* corresponds to the new components and *1s* corresponds to the remanufactured components. Initial recycled units (IRU) corresponds to the amount of remanufactured product that is supplied from node 1 to node 0. IRU can be considered as a *Min (demand for remanufactured components, available remanufactured components)*, in other words it is not always possible in every period to satisfy the entire amount of remanufactured component demand due to capacity constraints and yield at node 2, so only the available remanufactured components are supplied to node 1. These results in shortage of components to manufacture the final product, this deficit can be filled by new components if available in excess, after satisfying its corresponding portion of demand. This is referred to as ENU (excess new units). Similarly initial new units (INU) corresponds to the amount of new product that is supplied from node 1 to node 0. INU can be considered as a *Min (demand for new components, available new components)*, the amount of new components available in a given period may not always be able to satisfy the demand for the new products due to the unreliability in supply. In this situation only the available new product will be used to satisfy the component demand for the final product in node 0, it is possible that there could be excess unused remanufactured components available in same period. These excess unused remanufactured components can be used to satisfy the deficit caused due to new components. We refer this as ERU (excess remanufactured units). So the actual quantity of new components supplied from node 1 to node 0 is the sum of INU and ENU, and the actual quantity of remanufactured components from node 1 to node 0 is the sum of IRU and ERU.

The cost of the new components supplied by two of the supplier's changes dynamically in each period, depending on the cost offered by a supplier in a given period a decision is made. A default 40% of the new component demand has to be placed from a supplier irrespective of the cost. The lower cost supplier in a given period gets 60% of the new component demand, whereas the remainder 40% of the new component demand is procured from the high cost supplier.

The two-echelon problem formulation can be made as shown below:

$$\min_{s^i \geq 0} \sum_{i=0}^{1} c^i s^i + \overline{ENU} * c^{ENU} + \overline{ERU} * c^{ERU} + \varepsilon_n * IL_n^2$$
$$s.t\ P\left[IL_n^i \geq 0 \right] \geq \alpha^i, \text{where } i \in \{2, 1s, 1n, 0\} \tag{1}$$

Where $\overline{ENU}$, and $\overline{ERU}$ are the extra new units and extra remanufactured units averaged over the number of periods used in simulation. c^{ENU} and c^{ERU} is the cost for extra new units and extra remanufactured units respectively.

The outstanding order equations for this two echelon system for node 0, 1s, 1n, and 2 are described in equations (2-5) respectively, the outstanding orders for *1s* and *1n* is not constrained on the capacity, since we assume that there is no limit on the capacity of warehouse (node 1):

$$Y_{n+1}^0 = Y_n^0 + \xi - \min \left\{ \begin{array}{l} Y_n^0 + \xi, S^{1s} + S^{1n} - \left[Y_{n-2}^{1s} + Y_{n-2}^{1n} \right] - \xi_{n-1} - \xi_{n-2} \\ + \left[DS_{n-1}^{0s} + DS_{n-1}^{0n} + ERU + ENU \right], \eta_n^0 \end{array} \right\} \tag{2}$$

$$Y_{n+1}^2 = Y_n^2 + \left(\rho * \xi\right) - \min\left\{ Y_n^2 + \left(\rho * \xi\right), \varepsilon_n, \eta_n^2 \right\} \tag{3}$$

$$Y_{n+1}^{1s} = Y_n^{1s} + \rho * \xi - DS_n^{0s} - \min\left\{ Y_n^{1s} + \rho * \xi - DS_n^{0s}, IRP \right\} \tag{4}$$

$$Y_{n+1}^{1n} = Y_n^{1n} + (1-\rho) * \xi - DS_n^{0n} - \min\left\{ Y_n^{1n} + (1-\rho) * \xi - DS_n^{0n}, INP \right\} \tag{5}$$

On-hand inventory equations for Node 0, 2, 1n and 1s are described in equations (6-9) respectively:

$$I_n^0 = \max\left[0, S^0 - Y_{n-2}^0 - \left(\xi_{n-1} + \xi_{n-2} \right) + ERU + ENU \right] \tag{6}$$

$$I_n^2 = \max\left[0, S^2 - Y_{n-2}^2 - \left(\xi_{n-1} + \xi_{n-2} \right) * \rho \right] \tag{7}$$

$$I_n^{1s} = \max\left[0, S^{1s} - Y_{n-2}^{1s} - \left(\xi_{n-1} + \xi_{n-2} \right) * \rho + DS_{n-1}^{0s} \right] \tag{8}$$

$$I_n^{1n} = \max\left[0, S^{1n} - Y_{n-2}^{1n} - \left(\xi_{n-1} + \xi_{n-2} \right) * \phi + DS_{n-1}^{0n} \right] \tag{9}$$

Downstream shortages:

$$\begin{array}{l} DS_{n-1}^{0} = \max\left[0, \xi_{n-1} - \eta_{n-1}^{0} \right] \\ DS_{n-1}^{0s} = \rho * DS_{n-1}^{0} \\ DS_{n-1}^{0n} = \phi * DS_{n-1}^{0} \end{array} \tag{10}$$

4. Simulation and Computational Set-up

In this section we discuss briefly about the simulation, set-up for numerical results, different cases used, and few initial implications based on the results. ARENA is used to update the equations stated in the previous section, all the decision variables including the objective function and constraint is defined in OptQuest, which is a scatter search based optimization tool in ARENA. Upper and lower bound values of all the decision variables (base-stock levels, and ρ in few cases) is defined in OptQuest. A set of values for the decision variables is initially obtained from OptQuest, these decision variable values are used by ARENA and the simulation (updating the equations) run for 500 periods, at the end of simulation the service level value is fed as an input to OptQuest. Depending on the service level value a new set of values for decision variables is fed back to ARENA.

For all the computational results a three-period supply/manufacturing lead time is assumed. The values for demand, capacity, and yield are based on normal distribution. A service level of 90% is used for node 2 and 0. The following cost structure is considered for the numerical analysis: cost of remanufactured component < cost of ERU < cost new component < cost of ENU. For the numerical analysis, four instances each consisting of four scenarios are solved for each case. We consider 3 cases, and each case has several sub-cases. Table 1 lists the names of the instances and its associated coefficient of variation (CV) (a measure of variability, ratio of standard deviation and mean) for the demand and capacity. The capacity is denoted as "average" capacity when the mean capacity utilization is between 65% and 75%, and the capacity is defined as "tight" capacity when the mean capacity utilization is between 85% and 95%. The values of the mean demand is between 10 and 15. The variance for the demand and capacity is defined as high and low if the coefficient of variation is 0.4 and 0.2 respectively. Each instance consists of six scenarios (6 sets of values of demand and capacity).

Case 1, which can also be considered as a base case has four sub-cases, the sub-case is classified on the basis of yield for the remanufactured component. The four sub-cases consists of the following: i) a mean yield of 10 % with a CV of 0.2, ii) a mean yield of 25% with a CV of 0.4, iii) a mean yield of 10% with a CV of 0.4, and iv) a mean yield of 25% with a CV of 0.2. Each sub-case is run under all four instances. In case 2 we assume that we have perfect information about the yield, in other words we know about defective return products before they are purchased or accepted. So the yield for the remanufactured component is 0% or there is no yield. There are no sub-cases considered in case 2. For both the case 1 and 2 the supplier has an all-or-nothing supply, which means in a given period with a probability of *beta* a supplier can supply the entire order, or will supply nothing with a probability of *(1-beta)*. Case 3 has perfect information about the yield for remanufactured components, so there is no yield for the remanufactured component. Case 3 also has fraction of demand coming from the recycled products as a decision variable. Case 3 does not have an all-or-nothing supply from the supplier, but the supply is based on a uniform distribution. Three sub-cases are considered in case 2: i) supply distribution for supplier 1 = uniform (0.95, 1), supply distribution for supplier 2 = (0, 1) ii) supply distribution for supplier 1 = uniform (0.95, 1), supply distribution for supplier 2 = (0, 1), but we allocate 80 % of the new component demand to a supplier that states a lower cost among the two suppliers, and 20 % of the demand to the other supplier with higher cost, iii) supply distribution for supplier 1 = uniform (0, 1), supply distribution for supplier 2 = (0, 1).

Table 1: Instances for Two-echelon System

Instance #	Name of Instance	CV for Capacity	CV for Demand
1	Average Capacity (AC)	0.2	0.2
2	Tight Capacity (TC)	0.2	0.2
3	High Demand Variance with Average Capacity (HDVAC)	0.4	0.4
4	High Demand Variance with Tight Capacity (HDVTC)	0.4	0.4

We are still in process of compiling all the results and drawing conclusions, here are few important inferences just based of case 1: i) If the cost of the new product is slightly lower than the existing cost, then having a higher or a lower yield in presence of low variance in demand and capacity does not make a difference in total cost, ii) if the cost of the new product is higher than the existing cost, having a higher or a lower yield in presence of high variance in demand and capacity has a significant difference on total cost, iii) in case of a high variance in demand and capacity, we have lower price sensitivity on the total average cost, when compared to the low variance in demand and capacity, iv) for a higher yield situation, as the CV increases we find that the total cost reduces, due to the trade off in cost and the service level for the new product and recycled product. So it is always better to have more recycled product when yield is higher.

5. Conclusion

In this paper we study a two-echelon remanufacturing system under unreliable suppliers for new components. We develop update equations for the system and perform a simulation optimization under various scenarios. We find out what should be the optimal target stock levels for the remanufactured and new components, under realistic assumptions like stochastic demand, capacity, yield, arrivals and deterministic lead time. We describe the initial inferences based on the numerical analysis conducted for case 1. Specifically we try to find out best possible scenario that reduces the total supply chain cost under a given situation product mix, explore the behavior of the system under unreliable suppliers for new components, and optimal target stock level for new and remanufactured components.

6. References

[1] Bollapragada, R., U. S. Rao, J. Zhang. "Managing inventory and supply performance in assembly systems with random supply capacity and demand". *Management Science,* 50, 1729-1743.

[2] Fleischmann, M., van Nunen, J., & Gräve, B. "Integrating Closed-loop Supply Chains and Spare-Parts Management at IBM". *Interfaces*, 2003, 33, 44–56.

[3] Guide, V. D. R., & Van Wassenhove, L. N. "The Evolution of Closed-Loop Supply Chain Research". *Operations Research*, 2009, 57(1), 10-18.

[4] Kerr, W., & Ryan, C. "Eco-efficiency Gains from Remanufacturing a Case Study of Photocopier Remanufacturing at Fuji Xerox Australia". *Journal of Cleaner Production*, 2001, 75-81.

[5] Kumar, S., & Malegeant, P. "Strategic Alliance in a Closed-loop Supply Chain, a Case of Manufacturer and Eco-Non-Profit Organization". *Technovation*, 2006, 1127-1135.

[6] Lund, R. T. *The Remanufacturing Industry: Hidden Giant.* Boston, MA: Boston University

BUILDING AN ACCEPTANCE CHART AROUND THE PROCESS MEAN

Donald S. Holmes, Stochos Inc. P.O. Box 247, Duanesburg, N.Y. 12056.
(518) 895-2896, dsholmes@stochos.com
A. Erhan Mergen, Rochester Institute of Technology, Saunders College of Business
Decision Sciences, 107 Lomb Memorial Drive, Rochester, N.Y. 14623-5608.
(585) 475-6143, emergen@saunders.rit.edu

ABSTRACT

In this paper we will propose a new way of building an acceptance chart which could be utilized in Six Sigma environment. The proposed method is to build the acceptance chart around the desired value for the process mean rather than around the specification limits. An example is included to illustrate the proposed method.

Keywords: Acceptance charts, process control, process acceptance, Six Sigma program.

DISCUSSION

Process control charts, also known as Shewhart charts, deal with the process stability issue, i.e., they try to identify the types of variation (random (i.e., common) or special) that the process displays. A stable process is described as one that displays only a common cause of variation. Thus the question that is being tested on the control chart is "Is the process stable, i.e., in statistical control?" Quality practitioners like to keep their processes in stable condition so that they can predict their true capability and future performance and reduce variation and cost, etc., in the process. For more detailed information on control charts see, for example, Grant & Leavenworth [4], and Montgomery [8].

Acceptance charts, on the other hand, are designed to answer a different question. The question is not about the stability; it is about whether the process is producing parts that meet the customer specification limits. This question is very different from the one asked by the process control charts. Acceptance charts are suitable for processes, such as some chemical processes, that are expected to have unavoidable shifts in their average value but are still able to meet specifications set by the customer. In usual statistical process control terms, such a process is not considered in-control but may still be able to produce acceptable product. See, for example, Freund [3], Montgomery [8, pp.354-358], Holmes and Mergen [5, 6, 7] for more detailed explanations of the acceptance charts.

Acceptance charts are built around customer specification limits (one-sided or two-sided), using either 1) rejectable quality level and beta risk; 2) acceptable quality level and alpha risk; or 3) both rejectable and acceptable quality levels.

When the fluctuation of the process average is inevitable or allowed up to certain amount, acceptance charts are a good alternative, given that the process standard deviation is small enough compared to the width of the tolerance range, i.e., the difference between the upper and lower specification limits.

In this paper we will propose a new way of building an acceptance chart. The proposed method is to build the acceptance chart around the desired value for the process mean rather than around the specification limits; the desired value of the process mean could very well be the nominal value desired by the customer, if there is a nominal value. This proposed acceptance chart could be very useful, for example, in Six Sigma projects when the process mean is allowed to shift up to ± 1.5 standard deviations, i.e., it is assumed that the process average will not stay stable. In the Six Sigma environment, checking for the stability of the process mean may not be meaningful since the mean is allowed to shift. Thus the proposed acceptance chart would be a better alternative to check to see if the process average stays within the allowed range of fluctuation. The chart would help reduce the risk of stopping the process unnecessarily when the process mean stays within its allowed range of fluctuation. Use of conventional process control charts in this case, on the other hand, could give frequent and unnecessary "out-of-control" signals since they check a different question, i.e., "Has the process average shifted?" Thus, we are recommending the use of this proposed acceptance chart when the shift of the process mean is tolerated up to a certain level, such as the case in Six Sigma applications.

Statistically, Six Sigma means having no more than 3.4 defects per million opportunities in any process, product, or service. Pyzdek and Keller [9, page 3] describe Six Sigma as a program that incorporates elements from the work of many quality pioneers which aims for error-free business performance. The Six Sigma program involves the use of statistical tools within a structured methodology to gain the necessary knowledge to be more competitive by producing less expensive and better products/services faster and less expensively (Breyfogle [2] – also see Senturk, et al. [10] and Benitez et al. [1] for various Six Sigma applications). Those tools are applied through the DMAIC model: define the goals (D); measure the existing system (M); analyze the system to find ways to remove the gap between the current performance of the system and the desired performance (A); improve the system (I); and control (C) the new system (Pyzdek and Keller [9, pp.36-42]). The Six Sigma program, like other quality management programs, promotes being proactive rather than reactive. However, *"no more than 3.4 defects per million opportunities"* is valid under the assumption that the process average could shift up to ± 1.5 standard deviations, which yields about 4.5-sigma quality level. Six Sigma programs utilize many statistical, as well as some managerial tools. Based on our experience, however, acceptance charts are one of the least utilized tools in Six Sigma projects, despite the fact that conceptually these charts are very suitable for such projects given the assumption described above.

How to build the proposed acceptance chart:

The proposed acceptance charts will be based on a concept similar to the acceptable quality level, which will indicate how much fluctuation is desirable for the mean around the desired value, and the corresponding alpha risk. A constant, k_1, will determine this quantity, i.e., the amount of allowed fluctuation. In Six Sigma applications, since the process mean is allowed to shift up to ± 1.5 standard deviations, we will use $k_1=1.5$. Using $k_1=1.5$, the desired process average $(\bar{\bar{X}})$ and the process standard deviation (σ_X), the upper acceptable process mean (UAPM) value for the process mean will be calculated as follows (Figure 1):

$$UAPM = \bar{\bar{X}} + 1.5\sigma_X \tag{1}$$

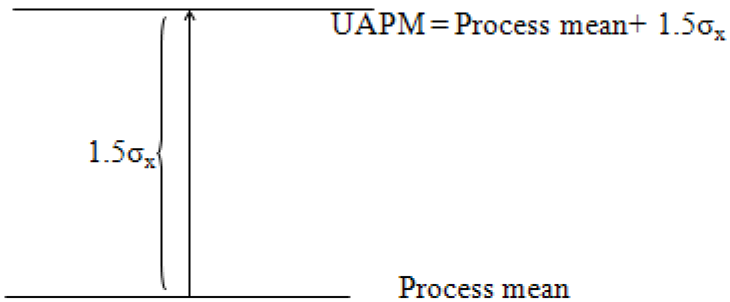


Figure 1. Graphical display of UAPM

This is the maximum value that we want to tolerate for the process mean. Then the upper acceptance limit (UAL) for the sample averages ($\bar{X}'s$) will be calculated as (Figure 2):

$$UAL = UAPM + k_2\sigma_{\bar{X}} \tag{2}$$

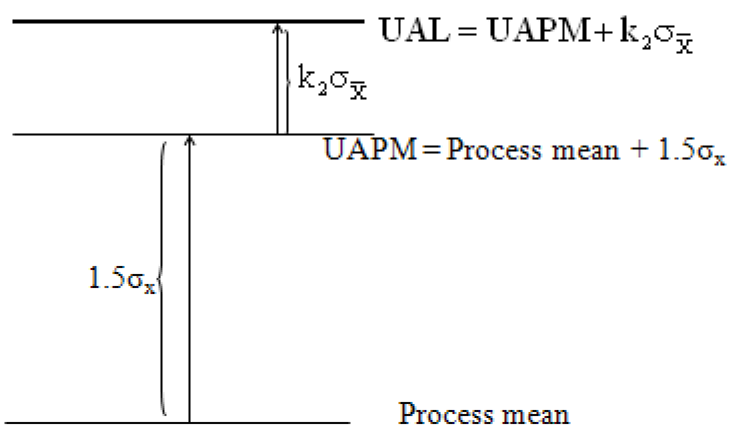


Figure 2. Graphical Display of UAL

where $\sigma_{\bar{X}}$ is the standard deviation of the sample averages ($\bar{X}'s$) and k_2 is a constant that would set the probability of accepting that the process average is not bigger than UAPM, which would also determine the resulting alpha risk (i.e., the type I error – rejecting the process when it is, in fact, running at the UAPM). In other words, we are addressing the question "Is the process behaving as the Six Sigma principle expects?" Remember that the standard deviation of the $\bar{X}'s$ is the standard deviation of the X's divided by the square root of the sample size, i.e., $\sigma_{\bar{X}} = \frac{\sigma_x}{\sqrt{n}}$ where n is the sample size.

Similar calculations could be done for the lowest acceptable process mean (LAPM) and the lower acceptance limit (LAL) as given below (Figure 3):

$$LAPM = \bar{\bar{X}} - 1.5\sigma_X \quad (3)$$

$$LAL = LAPM - k_2\sigma_{\bar{X}} \quad (4)$$

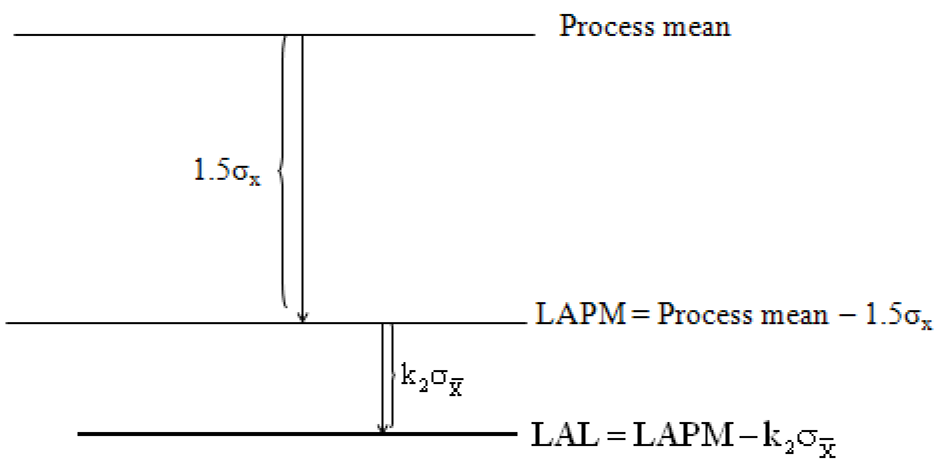


Figure 3. Graphical Display of LAPM and LAL

Thus, values of sample averages exceeding either UAL or LAL would be taken as a signal that the process average (mean) has shifted more than the allowed amount and that the proper corrective action should be taken. As long as the sample results fall between the acceptance limits, we will be accepting the fact that the process is behaving according to the Six Sigma model.

EXAMPLE

Let's assume that we have a process with a desired mean value of 50 and process output follows Normal distribution with standard deviation of 1. Let's also assume that process mean is allowed to fluctuate up to ±1.5 standard deviation around the desired value of 50. If we use k_2=3 and sample size (n) of 4, the acceptance limits for the proposed acceptance charts would be determined as follows:

$UAPM = 50 + 1.5(1) = 51.5$ $\qquad$ $UAL = 51.5 + 3\frac{1}{\sqrt{4}} = 53$

$LAPM = 50 - 1.5(1) = 48.5$ $\qquad$ $LAL = 48.5 - 3\frac{1}{\sqrt{4}} = 47$

Thus, as long as the sample averages fall between 47 and 53, we will accept (with 0.997 probability) that the process mean stays within ±1.5 standard deviation of the desired mean value of 50 (see also Figure 4 below). Thus no action would be required to adjust the mean.

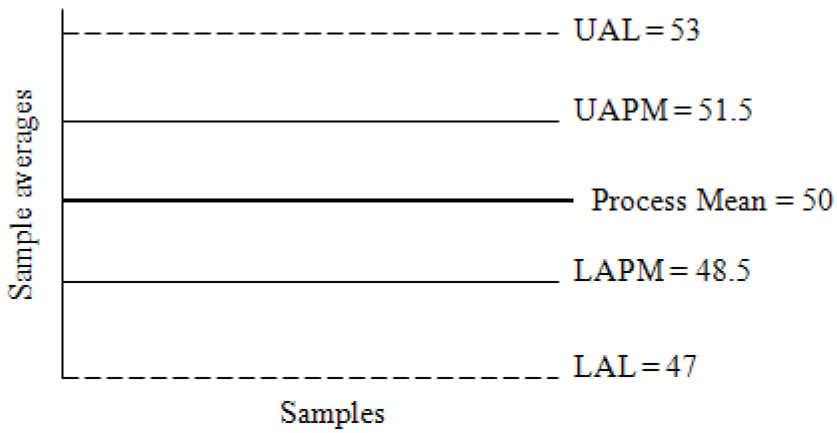


Figure 4. Acceptance chart around the mean.

CONCLUSION

In summary, in this paper we propose and describe a revised acceptance chart built around the process average to be used especially in Six Sigma projects when the process average is allowed to fluctuate to some degree. This proposed chart will help to keep the process average within the allowed range of fluctuation without increasing the alpha risk (i.e., the type I error). As a future study, a simulation analysis will be done to compare the false response rate of the control charts and the proposed acceptance chart when the shift in mean value is in the allowed range.

REFERENCES

[1] Benitez, Y., Forrester, L., Hurst, C. and Turpin, D. "Hospital reduces using DMAIC," Quality Progress, 2007, 40(1), 38-45.

[2] Breyfogle, F.W. *Implementing six sigma*. New York, NY: Wiley, 1999.

[3] Freund, R.A. "Acceptance control charts," *Industrial Quality Control*, October 1957, 14(4), 13-23.

[4] Grant, E.L. and Leavenworth, R.S. *Statistical quality control*. 6th edition, New York, NY: McGraw-Hill, 1988.

[5] Holmes, D.S. and Mergen, A.E. "Process acceptance charts for short runs," *Quality Engineering*, 1997, 10(1), 149-153.

[6] Holmes, D.S. and Mergen, A.E. "EWMA acceptance charts," *Quality and Reliability Engineering, International*, 2000, 16(2), 1-4.

[7] Holmes, D.S. and Mergen, A.E. "Use of Acceptance charts in the six sigma environment," *The 2009 Northeast Decision Sciences Institute Meeting Proceedings*, Uncasville, CT, April 1-3, 2009, 443-448.

[8] Montgomery, D.C. *Introduction to statistical quality control*. 4th edition, New York, NY: Wiley, 2001.

[9] Pyzdek, T. and Keller, P. *The six sigma handbook*. 3rd edition, New York, NY: McGraw-Hill, 2010.

[10] Senturk, D., LaComb, C., Neagu, R. and Doganaksoy, M. "Detect financial problems with six sigma," *Quality Progress*, 2006, 39(4), 41-47.

ON A SMALL OPTIMIZATION PROBLEM FOR A DISTRIBUTION CENTER

Hua Zhong
Economics & Business Division
State University of New York-Oneonta
Oneonta, NY 13820, USA
E-mail zhongh@oneonta.edu

Bin Zhou
College of Business and Public Administration
Kean University, Union, NJ 07083
E-mail bzhou@kean.edu

ABSTRACT

A common managerial decision problem encountered in daily operations of many distribution centers is the procurement of paper boxes for packing the products to customers. In this paper, we study the related packing box selection problem. The problem is to determine which box designs (and quantity) should be selected (and purchased) so that the total procurement cost is minimized. First, we use a greedy heuristic algorithm to select the box design. Then a linear programming model is used to determine the order quantity for each box design.

Key word: Distribution center; Greedy heuristic algorithm; Linear programming;

INTRODUCTION

A common managerial decision problem encountered in daily operations of many distribution centers is the procurement of paper boxes for packing the products to customers. Due to highly dynamic customer requirements on the variety of products which can be anything from the metal frame of computer hard disks to medical pills, the procurement decision on selecting packing boxes has been a major headache of the management. The main reason is that these paper boxes are usually ordered from external companies which charges a fixed cost for each box design and imposes a minimum order quantity for each selected box design. Such minimum ordering quantity often lead to a huge accumulation of unused boxes left from previous packing operations in manufacturer's inventory, and increase the waste and cost along supply chains.

In this paper, we study the related packing box selection problem. This problem can be stated as follows. We are given a set of products, each has its own size and quantity (in terms of the number of units) to be packed. Each unit of a given product requires the use of a packing box. We allow the same box design to be used for packing different products. However, additional filling material cost will incur if a box designed for a different (usually larger) product is used. For each selected box design, there is a minimum ordering quantity, a fixed overhead charge and a unit purchasing cost (i.e., the cost per box). The problem is to determine which box designs (and quantity) should be selected (and purchased) so that the total procurement cost is minimized.

MODEL FORMULATION

To make a formal presentation of our packing box selection problem, let

N The total number of product types (or the total number box designs);
d_j The quantity (in term of the number of boxes needed) of product j to be packed;
q_i The minimum order quantity for box design i, which fits exactly for product i,
w_i Inventory of box i at the beginning of the planning period;
f_i The fixed overhead cost for ordering box design i (or box i);
c_i The unit purchasing cost for box i;
e_{ij} Extra filling cost occurred when box i is used for packing product j. $e_{ii} = 0$,for all i, and $e_{ij} = +\infty$ if box i can not be used for product j.
Q_i The quantity of box design i, or box i, to be ordered.

$$z_i = \begin{cases} 1, & \text{if box } i \text{ is ordered} \\ 0, & \text{otherwise} \end{cases}$$

v_{ij} The total quantity of box i to be allocated for packing product j.
p_i The unit penalty cost for carrying each unit of box of design i to the next period.

Given the notations above, we can now formulate the packing box selection problem as the following mixed integer programming problem $\mathbf{P_0}$.

$$\mathbf{P_0}: \quad \text{Min } \{ TC = \sum_{i=1}^{N} (f_i z_i + c_i Q_i) + \sum_{j=1}^{N}\sum_{i=1}^{N} e_{ij} v_{ij} + \sum_{i=1}^{N} p_i (w_i + Q_i - \sum_{j=1}^{N} v_{ij}) \}$$

$$= \sum_{i=1}^{N} f_i z_i + \sum_{i=1}^{N} (c_i + p_i) Q_i + \sum_{j=1}^{N}\sum_{i=1}^{N} (e_{ij} + p_i) v_{ij} + \sum_{i=1}^{N} p_i w_i \tag{1}$$

s.t.

$$\sum_{j=1}^{N} v_{ij} \le Q_i + w_i, \quad i = 1, 2, \ldots, N, \tag{2}$$

$$\sum_{i=1}^{N} v_{ij} \ge d_j, \; j = 1, 2, \ldots, N, \tag{3}$$

$$Q_i \ge q_i z_i, \quad i = 1, 2, \ldots, N, \tag{4}$$

$$M z_i \ge Q_i, \quad i = 1, 2, \ldots, N, \tag{5}$$

$Q_i \ge 0$, $v_{ij} \ge 0$, $w_i \ge 0$ and $z_j = 0$, or 1.

where constraints (2) ensure that the total usage of box design i does not exceeds the total quantity available. Constraints (3) ensure all the units of product j are packed in boxes. Constraints (4) ensure the minimum ordering quantity requirements are satisfied, and constraints (5) are used to make sure that the values of z_i are correctly assigned.

The above formulation allows more than one type of boxes to be used for packing a given product. While such flexibility may lead to a lower total cost, it may not be acceptable in reality since it causes confusion in packing operations and to the customers. In real life distribution centers, it is a common practice to use only the same type of boxes for the same product. The following property shows that if we relax he minimum order quantity requirement (i.e. q_i = 0, for all i), such a consideration can be met automatically.

***Property 1*:** Assume that the minimal box order quantity $q_i = 0$, for all i. Then, there is at least one optimal solution of the problem $\{ z_j, v_{ij}\}$, such that, for any given product j, there is only one i^o such that $v_{i^o j} = d_{i^o}$, and $v_{ij} = 0$ for all other i (proof given in the working paper).

Mathematically, if we assume that only one type of box should be used for packaging each product, then, the definition v_{ij} in the above formulation should be changed to $u_{ij}d_j$ where

$$u_{ij} = \begin{cases} 1, & \text{if box } i \text{ is used for product } j \\ 0, & \text{otherwise} \end{cases}$$

And, equation (3) should be changed to

$$\sum_{i=1}^{N} u_{ij} = 1, \quad j = 1, 2, \ldots, N, \qquad (3')$$

ALGORITHM OUTLINE

The following algorithm uses a greedy heuristic algorithm to select the box designs. After the box designs are fixed, $\mathbf{P_0}$ reduces to a linear programming problem and can be solved easily to further determining the order quantity for each selected box design to meet the packing needs.

Let's use i>j to denote that the box of product i can be used to pack product j. If i>j and j>i, then we keep the partial order that minimizes the filling cost. Leti i<>j to denote that box i can not be used for packing product j and box j can not be used for packing product i. In addition, we introduce with following notation:

$P_0(i)$ = The set of products that can be packed by using box i, which satisfies (a) if $k \in P_0(i)$, then $i > k$; and (b) if both $k_1 \in P_0(i)$, $k_2 \in P_0(i)$, $k_1 \neq k_2$, then $k_1 <> k_2$.

$B_0(j)$ = The set of boxes that can be used to pack product j. However, if both $k_1 \in B_0(i)$, $k_2 \in B_0(i)$, $k_1 \neq k_2$, then $k_1 <> k_2$.

$P_1(i)$ = The set of products that is currently assigned to box i.

$b(j)$ = The box design that is currently assigned to pack product j.

From the definition above, we know that for box i, if $P_1(i) \neq \emptyset$, then, box i will be ordered and the corresponding costs relate to products in $P_1(i)$ is

$$C(P_1(i)) = f_i + c_i\, x(P_1(i)) + y(P_1(i)), \text{ where}$$

$x(P_1(i)) = \max\{\sum_{j\in P_1(i)} d_j, q_i\}$ defines the quantity of box i to be ordered, and $y(P_1(i)) = \sum_{j\in P_1(i)} e_{ij} d_j$ stands for the total filling cost of using box i to pack products in $P_1(i)$.

PSEUDO CODE OF THE ALGORITHM

(Input: $P_0(i)$, $B_0(i)$, and a partial order of product/box designs, <1, 2, …, N>, where i>i+1, for all i.

***Step* 0.** Initialization:
Let $n = 1$, $S = \{1, 2, \ldots, n\}$, $B_1(1) = 1$, $P_1(1) = \{1\}$, let the total cost associated with products in S be

$$C(S) = \sum_{j\in S} [f_j + c_j\, x(P_1(i))] = \sum_{j\in S} [f_j + c_j\, (\max\{d_j, q_j\})] \tag{6}$$

***Step* 1.** Consider the following two cases:
Case a): If product n has its own box selected, then the corresponding total costs after adding the product to S becomes

$$C_a(S\cup\{n\}) = \mathrm{C}(S) + f_n + c_n\, x(\{n\}) \tag{7}$$

Case b). If product n does not have its own box, then for each $k \in B_0(n)$:

*Case b*1. $P_1(k) \neq \varnothing$, i.e. box k has been selected. If we also use box k to pack product n, the corresponding cost becomes

$$C_{b1}(S\cup\{n\}, k) = C(S) - c_k\, x(P_1(k)) + c_k\, x(P_1(k)\cup\{n\}) + e_{kn}\, d_n \tag{8}$$

*Case b*2. $P_1(k) = \varnothing$, then, $b(k) = h \neq k$.

Case b21. If box h is also to be used to pack product n, the corresponding cost will be

$$C_{b21}(S \cup\{n\}, k) = \mathrm{C}(S) - c_h\, x(P_1(h)) + [c_h\, x(P_1(h)\cup\{n\})] + e_{hn}\, d_n, \text{ or,} \tag{9}$$

*Case b*22. If box k will be ordered to pack both product k and product n, then the corresponding cost becomes

$$C_{b22}(S \cup\{n\}, k) = \mathrm{C}(S) + c_h\, [x(P_1(h)\backslash\{k\}) - x(P_1(h))] + f_k + c_k\, [x(\{k\}\cup\{n\})] \tag{10}$$

Let $C(S \cup\{n\}|\, k^*)$ be the minimum of (8), (9) and (10), over all $k\in B_0(n)$. Then, let
$S \leftarrow S \cup\{n\}$, and
$b(n) = n$, $P_1(n) = \{n\}$, if *(7)*< $C(S \cup\{n\}|\, k^*)$;
$b(n) = k$, $P_1(n) = \{\varnothing\}$, and $P_1(k) \leftarrow P_1(k) \cup\{n\}$, if $C(S \cup\{n\}|\, k^*)$ occurs at (8);
$b(n) = h$, $P_1(n) = \{\varnothing\}$, and $P_1(h) \leftarrow P_1(h)\cup\{n\}$, if $C(S \cup\{n\}|\, k^*)$ occurs at (9);
$b(n) = k$, $P_1(n) = \{\varnothing\}$, and $B_1(k) = k$, $P_1(h) \leftarrow P_1(h)\backslash\{k\}$, $P_1(k)\leftarrow\{k, n\}$, if $C(S \cup\{n\}|\, k^*)$ occurs at (10).
Goto ***Step* 1.**

***Step* 2.** Let $z_j = 1$, if $P_1(j) \neq \varnothing$, and $z_j = 0$, if $P_1(j) = \varnothing$.

It can be proved that the algorithm developed above has a computation complexity less than $O(N^2)$, where N is the total number of product types. Also, as we can see from model $\mathbf{P_0}$, if z_is are fixed, then $\mathbf{P_0}$ reduces to a linear programming model which can be solved easily even with a very large N.

EMPIRICAL PERFORMANCE

To observe the performance of the proposed algorithm in a more general setting, we randomly generated test cases based on some hypothetical assumptions. A few sets of experiments, each investigates the impact of a particular problem parameter on the computational performance of the proposed algorithm, were conducted (result given in the working paper). As we can see, the proposed algorithm showed a very consistent and promising empirical performance in terms of both CPU time required and its deviations from the optimal solutions.

REFERENCES

[1] Ahuja, R. K., T.L. Magnanti & J. B. Orlin, 1993. *Network Flows*. Prentice Hall.
[2] Armstrong, R., L. Lei and S. Gu, 1998. "An Efficient Algorithm for a Class of Two-resource Allocation Problems". *INFORMS Journal on Computing***.** Vol. 10, No. 1, pp. 114-120.
[3] Bazaraa, M., J. Jarvis and H. Sherali, 1990. *Linear Programming and Network Flows*. Wiley.
[4] Kuno, T., H. Konno and E. Zemel, 1991. "A Linear-time Algorithm for Solving Continuous Maximum Knapsack Problem". *Operations Research Letters*. Vol. 10, pp. 23-26.

AN OPTIMAL PRODUCTION CONTROL PROBLEM

G. Naadimuthu, Fairleigh Dickinson University, 973-443-8850, naadi@fdu.edu
S. Waziruddin, Kansas State University, 785-532-5606, eslee@ksu.edu
E. S. Lee, Kansas State University, 785-532-5606, eslee@ksu.edu

ABSTRACT

A constrained production planning problem with a known sales rate is considered. The objective is to determine the production rate that minimizes a cost function involving production and inventory. A first variational technique is shown to be an effective approach for this optimization problem. Convergence rates of production, inventory, and cost are obtained.
Keywords: Production control; first variational technique; optimization

A PRODUCTION CONTROL PROBLEM

There have been several publications [1-3, 5-8] in the area of production and inventory control. This paper is distinct since it deals with the application of a first variational technique in the optimization of a production and inventory control problem described by differential equations with an integral objective function.

Let us consider the following production planning problem.

Assume that the rate of sales at time t, Q (t), is given by

$$Q(t) = a + bt \tag{1}$$

where a and b are known constants. The problem is to determine the value of the decision variable, namely the production rate (p(t)) subject to the following rate of change in inventory level (I(t)) constraint

$$\frac{dI(t)}{dt} = p(t) - Q(t), \quad 0 \le t \le T \tag{2a}$$

$$I(0) = I^0 \tag{2b}$$

$$I(T) = C \tag{2c}$$

so that the cost function

$$J = \int_0^T \left(C_I I^2(t) + C_p \exp[p_m - p(t)]^2 \right) dt \tag{3}$$

is minimized. J is the total cost of production and inventory. C_I is the unit inventory carrying cost. C_p is the minimum production cost which occurs when production rate equals p_m. p_m can be considered as the normal capacity of the manufacturing plant. Any fluctuation of production rate from p_m incurs additional cost due to either over or under usage of available manpower and equipment.

To solve the above problem, let us introduce an additional state variable as follows.

$$x(t) = \int_0^t (C_I I^2(t) + C_P \exp[p_m - p(t)]^2)dt \tag{4}$$

Equation (4) can be rewritten as

$$\frac{dx(t)}{dt} = C_I I^2(t) + C_P \exp[p_m - p(t)]^2 \tag{5a}$$

$$x(0) = 0 \tag{5b}$$

The problem now reduces to seeking the value of the decision variable p(t) so that the state variable x(t) is minimized.

Following the treatment of a first variational technique [4], the recurrence equations for the production planning problem are

$$\frac{\partial S_1}{\partial p}\Big|_t = \frac{\partial S_1}{\partial I}\Big|_{t+\Delta}\Delta - \frac{\partial S_1}{\partial x}\Big|_{t+\Delta}\left(2C_p\left(p_m - p(t)\right)\exp\left(p_m - p(t)\right)^2\right)\Delta \tag{6}$$

$$\frac{\partial S_1}{\partial I}\Big|_t = \frac{\partial S_1}{\partial I}\Big|_{t+\Delta} + 2\frac{\partial S_1}{\partial x}\Big|_{t+\Delta} C_I I\Big|_t \Delta \tag{7}$$

$$\frac{\partial S_1}{\partial x}\Big|_t = \frac{\partial S_1}{\partial x}\Big|_{t+\Delta} \tag{8}$$

The final conditions are

$$\frac{\partial S_1}{\partial I}\Big|_T = 0 \tag{9}$$

$$\frac{\partial S_1}{\partial x}\Big|_T = 1 \tag{10}$$

The given final condition (2c) can be rewritten as

$$\xi = I(T) - C = 0$$

In order to satisfy this final condition, we have

$$\frac{\partial S_2}{\partial p}\Big|_t = \frac{\partial S_2}{\partial I}\Big|_{t+\Delta}\Delta - 2\frac{\partial S_2}{\partial x}\Big|_{t+\Delta} C_p\left(p_m - p(t)\right)\exp\left(p_m - p(t)\right)^2\Big|_t \Delta \tag{11}$$

$$\frac{\partial S_2}{\partial I}\Big|_t = \frac{\partial S_2}{\partial I}\Big|_{t+\Delta} + 2\frac{\partial S_2}{\partial x}\Big|_{t+\Delta} C_I I\Big|_t \Delta \tag{12}$$

$$\frac{\partial S_2}{\partial x}\Big|_t = \frac{\partial S_2}{\partial x}\Big|_{t+\Delta} \tag{13}$$

$$\frac{\partial S_2}{\partial I}\Big|_T = 1 \tag{14}$$

$$\frac{\partial S_2}{\partial x}\Big|_T = 0 \tag{15}$$

The numerical values used in solving the problem are

$a = 2.0$	$C = 9.25$
$b = 1.0$	$C_p = 0.001$
$C_I = 0.1$	$p_m = 5.0$
$T = 1.0$	$\Delta = 0.01$
$I^0 = 5.0$	

The convergence rate of the decision variable, production, using an initial approximation of p(0) = 7.0 is shown in Figure 1. Also, the convergence rates of the state variables, inventory and cost, are displayed in Figures 2 and 3 respectively. The Runge-Kutta integration formula was used to integrate Equations (2) and (5). The step size used was 0.01 which is the same as the Δ value used. A value of $\Delta \phi$ equal to - 0.1 was used for the first 25 iterations and values of $\Delta \phi$ = - 0.01 and - 0.001 were used for 26 to 72 and 73 to 126 iterations, respectively. The required end condition on inventory was satisfied after two iterations. The convergence rate of the cost J was extremely slow after the fifth iteration. The cost at the fifth iteration was 5.25. It took approximately 90 iterations to improve this value from 5.25 to the minimum of 5.17.

REFERENCES

[1] Dennis, D. R. and Meredith, J. R. “An Analysis of Process Industry Production and Inventory Management Systems.” *J. of Operations Management*, 2000, 18, 683.

[2] Disney, S. M. and Grubbstrom, R. W. “Economic Consequences of a Production and Inventory Control Policy.” *Int. J. of Production Research*, 2004, 42, 3419.

[3] He, Q. M., Jewkes, E. M., and Buzacott, J. “Optimal and Near-Optimal Inventory Control Policies for a Make-to-Order Inventory-Production System.” *European J. of Operational Research*, 2002, 141, 113.

[4] Lee, E, S. "Optimization by a Gradient Technique." *I & EC Fundamentals*, 1964, 3, 373.

[5] MacCarthy, B. L. and Fernandes, F. C. F. “A Multi-Dimensional Classification of Production Systems for the Design and Selection of Production Planning and Control Systems.” *Production Planning Control*, 2000, 11, 481.

[6] Minner, S. "Multiple-Supplier Inventory Models in Supply Chain Management: A Review." *Int. J. of Production Economics*, 2003, 81-82, 265.

[7] Riddalls, C. E. and Bennett, S. "Production-Inventory System Controller Design and Supply Chain Dynamics." *Int. J. of Systems Science*, 2002, 33, 181.

[8] Yang, P. C. and Wee, H. M. "An Integrated Multi-Lot-Size Production Inventory Model for Deteriorating Items." *Computers and Operations Research*, 2003, 30, 671.

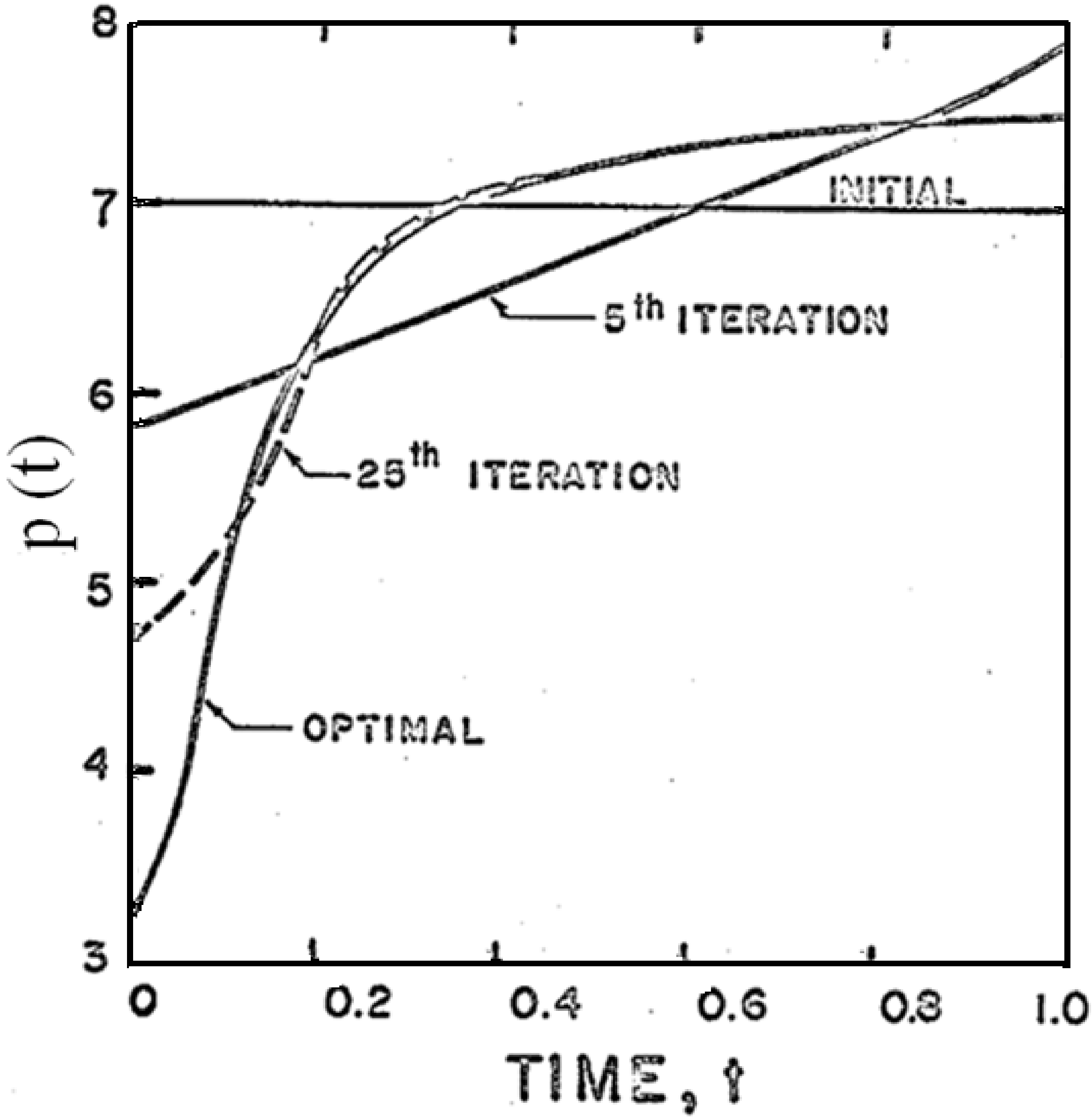


FIGURE 1. CONVERGENCE RATE OF PRODUCTION

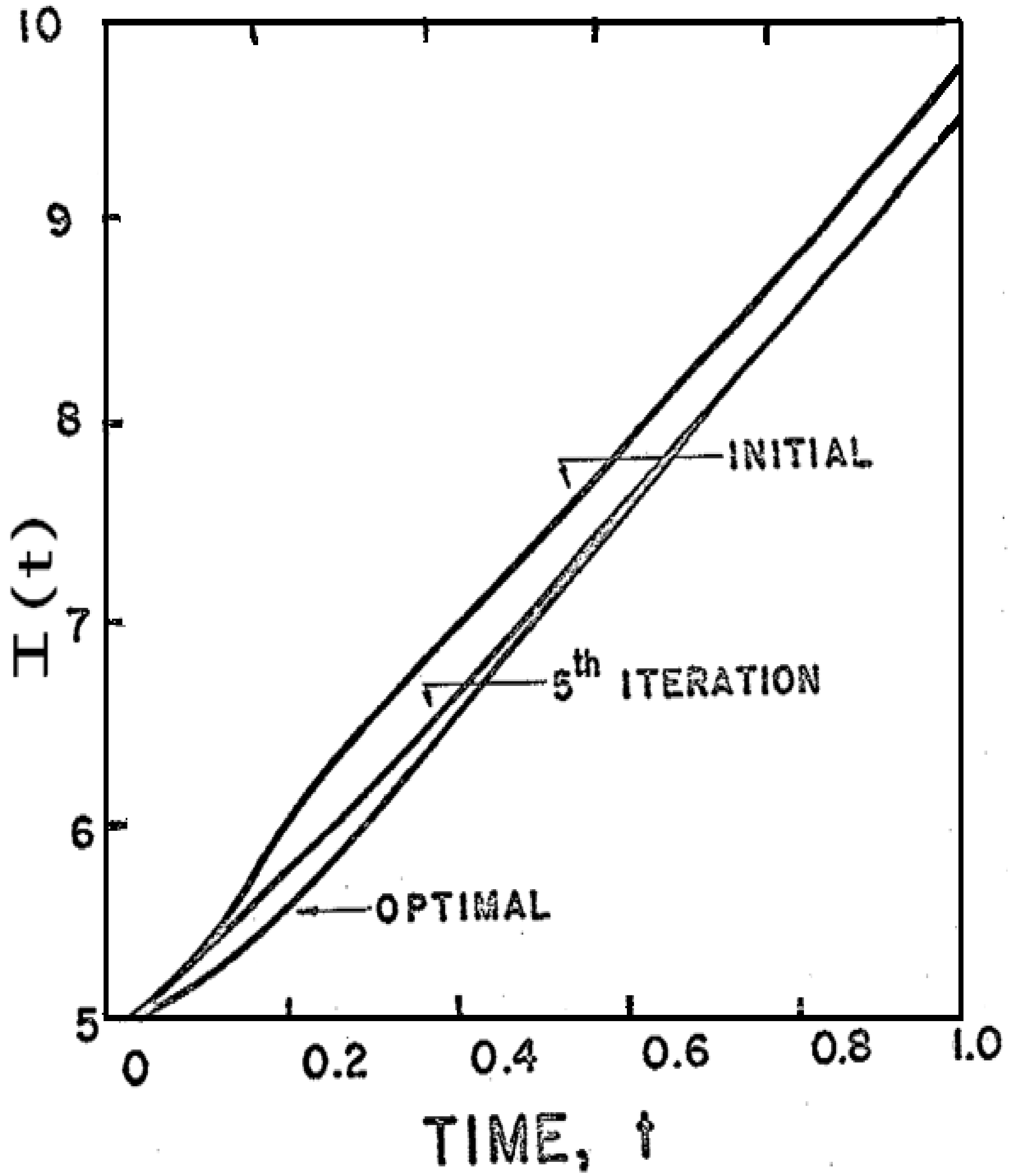


FIGURE 2. CONVERGENCE RATE OF INVENTORY

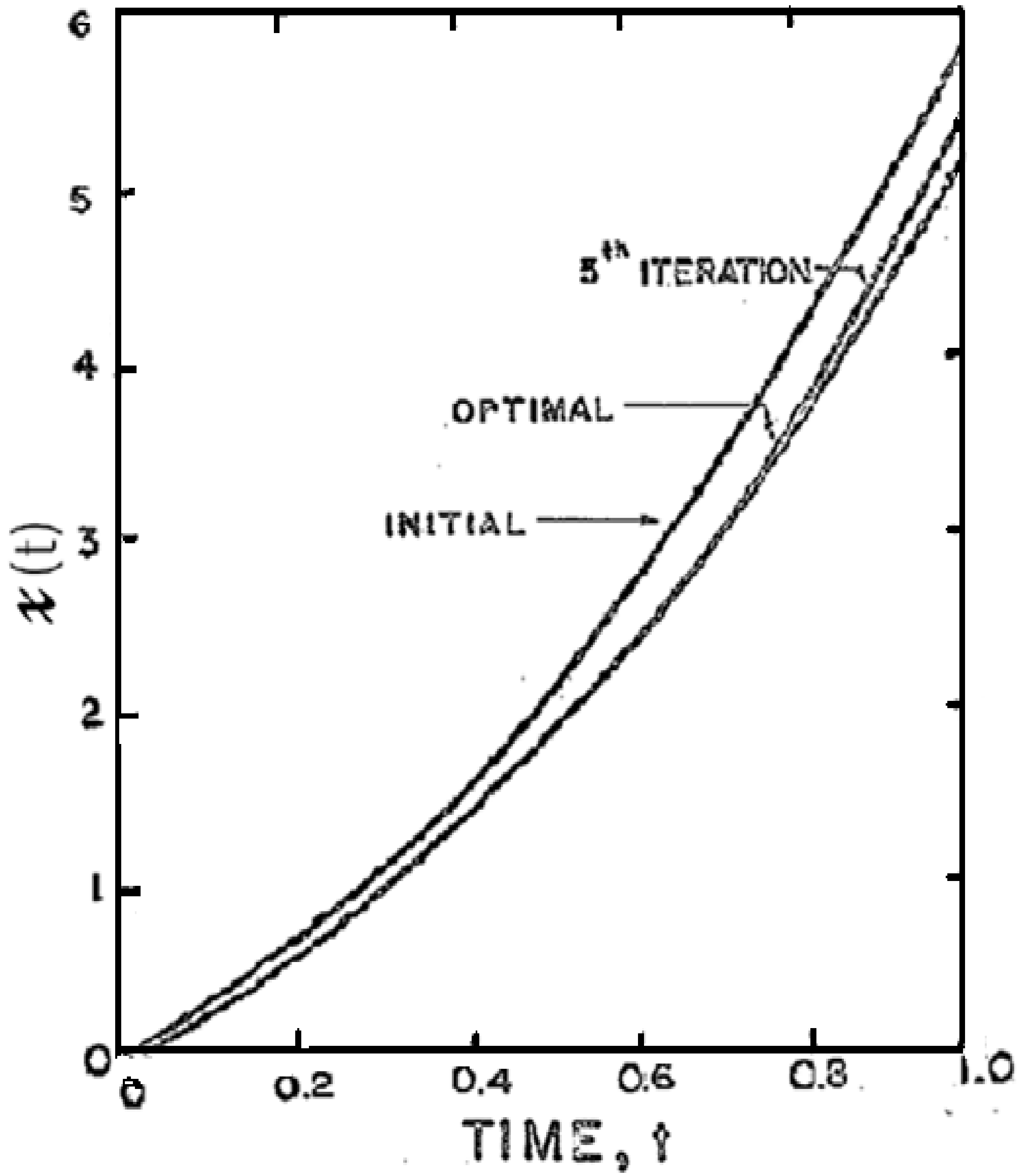


FIGURE 3. CONVERGENCE RATE OF COST

THREE HEURISTSICS FOR COORDINATED PLANNING OF A PRODUCT CONTROLLED BY THE EPQ WITH PARTIAL BACKORDERING AND ITS COMPONENTS

David W. Pentico, School of Business, Duquesne University, Pittsburgh, PA 15282-0180, pentico@duq.edu, 412-396-6252
Matthew J. Drake, School of Business, Duquesne University, Pittsburgh, PA 15282-0180, drake987@duq.edu, 412-396-1959
Carl Toews, Department of Mathematics, Duquesne University, Pittsburgh, PA 15282-1754, toewsc@duq.edu, 412-396-4851

ABSTRACT

We complement a previously developed optimization model that recognizes the WIP effects when there is gradual conversion of components into a final product by considering three heuristics for lot-sizing planning for a two-stage system in which the final product is planned using an EPQ model with partial backordering and the production of the components is controlled using basic EPQ models without backordering.

1. INTRODUCTION

We extend previous work on an optimization model discussed in Pentico et al. (2009b) by considering three heuristics for the coordinated planning of a two-level system consisting of a final product, which has the characteristics of an EPQ with partial backordering, and its immediate components, which are subject to control by an EPQ model without backordering.

2. THE BASIC MODEL

For the final product, we make all the usual assumptions of a deterministic EPQ model with full backordering except we assume that only a constant fraction β of the demands that occur during the stockout period will be backordered. In addition, we make the following assumptions:

- FIFO backorder filling for the final product, as discussed in Pentico et al. (2009a), meaning that the oldest backorders are filled first;
- a one-stage final production system, as in Banerjee et al. (1990), Sharma and Singh (2004) and Pentico et al. (2009b);
- the final product has m components (we assume that there is one unit of each component by redefining a component "unit" to consist of multiple original units, adjusting the holding cost and production rate for the component to compensate for the redefinition);
- perfect quality of the components and the final product;
- the production of components is according to an EPQ process with no backordering.

Parameters

D = demand per year for the final product
P = production rate per year of the final product if constantly producing
P_i = production rate per year of component i if constantly producing ($P_i > P$)

s = the unit selling price for the final product
C_o = the fixed cost of placing and receiving an order for the final product
C_{oi} = the fixed cost of placing and receiving an order for component i
C_p = the variable cost of producing a unit of the final product
C_{pi} = the variable cost of producing a unit of component i (included in the final product's cost C_p)
C_h = the cost to hold a unit of the final product in inventory for a year
C_{hi} = the cost to hold a unit of component i in inventory for a year
C_b = the cost to keep a unit of the final product backordered for a year
C_g = the goodwill loss on a unit of unfilled demand for the final product
$C_l = (s - C_p) + C_g$ = the cost of a lost sale for the final product, including the lost profit on that unit and any goodwill loss
β = the fraction of stockouts of the final product that will be backordered

Variables

Q = the order quantity for the final product
Q_i = the order quantity for component i
T = the length of an order cycle for the final product
F = the fill rate or the percentage of demand for the final product that will be filled from stock
N_i = the number of times component i is produced/ordered during a production interval for the final product (an integer)

The objective function, as shown in Pentico et al. (2009b), combines the average cost per year for the final product with the average cost per year for the components:

$$\Gamma(T,F,N_1,\ldots,N_m) = \frac{C_o + \sum_{i=1}^{m} N_i C_{oi}}{T} + C_h' \frac{DTF^2}{2} + \frac{TD^2[\beta+(1-\beta)F]^2}{2P} \sum_{i=1}^{m} \frac{C_{hi}'}{N_i} + C_b' \frac{\beta DT(1\text{-}F)^2}{2} + C_l D(1-\beta)(1-F) \quad (1)$$

where $C_h' = C_h(1 - D/P)$, $C_b' = C_b(1 - \beta D/P)$, and $C_{hi}' = C_{hi}(1 - P/P_i)$.

3. THE OPTIMIZATION MODEL

As described in Pentico et al. (2009b), the optimization is based on determining values for T, F, and continuous values for the N_is by using classical calculus-based optimization methods and then integerizing the N_is, using a method described in Banerjee et al. (1990). The continuous values for the N_is are given by:

$$N_i^* = \frac{DT[\beta+(1-\beta)F]}{\sqrt{2P}} \sqrt{\frac{C_{hi}'}{C_{oi}}}. \quad (2)$$

Equation (2) is then substituted back into (1) to give a cost function based on T and F, from which equations for T^* and F^* and conditions under which they apply are developed.

Banerjee et al.'s (1990) procedure for integerizing the N_is is as follows: Fix any N_is with integer values at those values. (If any N_i is less than 1, its integerized value must be 1.) Beginning with the non-integer N_i with the lowest index, let $\underline{N_i}$ be the largest integer less than N_i and let $\overline{N_i}$ be the smallest integer greater than N_i. ($\underline{N_i}$ and $\overline{N_i}$ "surround" N_i in Banerjee et al.'s terminology.) Substitute $\underline{N_i}$ and $\overline{N_i}$, along with any N_is already fixed at integer values and the remaining non-integer N_is into the objective function. Fix the value of N_i for this index i to be the surrounding value that gives the lower objective function value. Repeat until all components have had their value of N_i fixed as an integer.

4. THREE HEURISTICS

Although finding the "optimal" solution is not difficult, there are other solution methods that are either easier to implement or have other desirable characteristics. In this section we consider three heuristics: A) use Pentico et al.'s (2009a) optimal solution procedure for the EPQ with partial backordering and no components (EPQ-PBO) to determine the values of T and F, and then use (2) to determine the initial values of the N_is, which will then be integerized; B) use the same integerized value of N for all components; and C) use lot-for-lot (L4L) ordering of the components (i.e., let $N = 1$ for all components).

4.1 Heuristic A: Use Pentico et al.'s (2009a) EPQ-PBO model for T and F

Pentico et al.'s (2009a) model controls only the production of the final product, so it does not include equations for the N_is. This heuristic uses Pentico et al.'s (2009a) equations for T and F and then uses (2) to get the optimal values of the N_is for that T and F, which are then integerized.

Step 1: Use the EOQ-PBO equations from (2009a) to get values for T and F:

$$T_A^* = \sqrt{\frac{2C_o}{DC_h'}\left[\frac{C_h' + \beta C_b'}{\beta C_b'}\right] - \frac{[(1-\beta)C_l]^2}{\beta C_h' C_b'}} \qquad F_A^* = F(T_A^*) = \frac{(1-\beta)C_l + \beta C_b' T_A^*}{T_A^*(C_h' + \beta C_b')} \tag{3}$$

Step 2: Use (2) to find the optimal continuous values for the N_is given T and F.

Step 3: Integerize the N_is.

4.2 Heuristic B: Use the same N for all components

The concept of this heuristic is to fully coordinate the production of the components by ordering them all at the same time. This heuristic may be particularly of interest if all the components are purchased from the same supplier so that they may be shipped in together. (Using the EOQ assumptions instead for the components requires simply replacing C_{hi}' by C_{hi} in the equations.) Dropping the subscripts on the Ns in Eq. (1) gives the cost equation for Heuristic B, from which the equation for the optimal continuous value of N can be determined to be:

$$N = \frac{DT[\beta + (1-\beta)F]}{\sqrt{2P}} \sqrt{\sum_{i=1}^{m} C_{hi}' \Big/ \sum_{i=1}^{m} C_{oi}} \tag{4}$$

The optimal values of T and F given the requirement that all components use the same N are then found using the same approach as in the full optimization model given in Pentico et al. (2009b).

Step 1: Use the following equations to find the values of G_0, G_1, G_2, and G_3.

$$G_0 = C_0 \qquad G_1 = \frac{D(C_h^{'} + \beta C_b^{'})}{2} \qquad G_2 = \frac{\beta C_b^{'} D}{2} \tag{5}$$

$$G_3 = \frac{2D(1-\beta)\sqrt{\sum_1^m C_{hi}^{'} \sum_1^m C_{oi}}}{\sqrt{2P}} - C_l D(1-\beta) \qquad G_4 = \frac{2D\beta\sqrt{\sum_1^m C_{hi}^{'} \sum_1^m C_{oi}}}{\sqrt{2P}} + C_l D(1-\beta)$$

Step 2: If $G_3 - 2\sqrt{G_0 G_2} > 0$, set $T = \infty$ and $F = 0$. The product should not be produced. If not, go to Step 3.

Step 3: If $G_3 + 2\sqrt{G_0(G_1 - G_2)} < 0$, set $F = 1$ and $T = \sqrt{2C_o/(DC_h^{'})}$. The product should be produced with no backordering. Go to Step 7. If not, go to Step 4.

Step 4: Compute the values of a, b, and c as:

$$a = G_1 G_3^2 - 4G_0 G_1^2 \qquad b = 8G_0 G_1 G_2 - 2G_2 G_3^2 \qquad c = G_2 G_3^2 - 4G_0 G_2^2$$

Step 5: Use the quadratic formula to determine two values of F:

$$F_1, F_2 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Step 6: Use F_1 and F_2 to determine two values of T from:

$$T = -G_3/(2(G_1 F - G_2))$$

Set $T^* = T_1$ or T_2, whichever is positive. Set F^* = the value of F that gave T^*.

Step 7: Use (4) to determine the continuous value of N.
Step 8: Integerize N.

4.3 Heuristic C: Lot-for-lot

Although this may not be the best JIT strategy, we shall consider it since L4L is mentioned fairly often in the literature as a possible method for scheduling the production of components for a finished product. We start by noting that this is a special case of Heuristic B in which $N = 1$ for all components. Replacing N by 1 in the cost function for Heuristic B gives a cost equation in T and F only which can be optimized as follows:

Step 1: Use the following equations to compute the values of G_0, G_1, G_2, G_3, and G_4.

$$G_0 = C_0 + \sum_{i=1}^m C_{oi} \qquad G_1 = \frac{D(C_h^{'} + \beta C_b^{'})}{2} + \frac{D^2(1-\beta)^2 \sum_{i=1}^m C_{hi}^{'}}{2P} \tag{6}$$

$$G_2 = \frac{2\beta(1-\beta)D^2 \sum_{i=1}^m C_{hi}^{'}}{2P} - \frac{2\beta C_b^{'} D}{2} \qquad G_3 = \frac{\beta C_b^{'} D}{2} + \frac{D^2 \beta^2 \sum_{i=1}^m C_{hi}^{'}}{2P} \qquad G_4 = -C_l D(1-\beta)$$

Step 2: If $G_2\sqrt{G_0/G_3} + G_4 > 0$, set $T = \infty$ and $F = 0$. The product should not be produced. If not, go to Step 3.

Step 3: If $\frac{\sqrt{G_0}(2G_1+G_2)}{\sqrt{G_1+G_2+G_3}}+G_4 < 0$, set $F = 1$ and $T = \sqrt{2C_o/(DC_h^{'})}$. The product should be produced with no backordering. If not, go to Step 4.

Step 4: Compute the values of a, b, and c as:

$$a = G_1G_4^2 - 4G_0G_1^2 \qquad b = G_2G_4^2 - 4G_0G_1G_2 \qquad c = G_3G_4^2 - G_0G_2^2$$

Step 5: Use the quadratic formula to determine two values of F:

$$F_1, F_2 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Step 6: Use F_1 and F_2 to determine two values of T from:

$$T = -G_4/(2G_1F + G_2)$$

Set $T_C^* = T_1$ or T_2, whichever is positive. Set F_C^* = the value of F that gave T_C^*.

5. Computational Study

To gain some insights into how well these three heuristics work, we identified six factors that we thought might make a difference in performance: number of components = 2 or 8; ratio of C_o to C_h for the final product = 0.5 or 5.0; ratio of C_l to C_b for the final product = 1.25 or 2.50; ratio of P to D for the final product = 3 or 10; value of β = 0.75 or 0.95; and the mix of desired ordering frequencies for the components is all monthly, all weekly, or half monthly and half weekly. The combinations gave a total of 96 cases.

All 96 cases were solved with the continuous optimization procedure in Pentico et al. (2009b), which gives a lower bound on the optimal cost. Each problem was also solved using the three heuristics described in Section 3. The performance measure used is the ratio of a heuristic's cost to that lower bound.

Heuristic A did extremely well, with an average cost ratio of 1.00001, a maximum cost ratio of 1.0001, and an essentially optimal solution in 93 of the 96 cases.

Somewhat surprisingly, Heuristic B – finding the optimal solution with the added requirement that all components use the same integerized N – did well. While it found an essentially optimum solution in only 39 of the 96 cases, the average cost ratio was less than 1.001, or less than 0.1 percent above the optimum, and the maximum ratio was 1.0069. Although we cannot generalize the results of this study to the broader set of possible applications of these rules, they do suggest that fully coordinating the production or receipt of components should increase costs very little.

As might be expected, Heuristic C – ordering components on a lot-for-lot basis relative to the orders for the final product – did not do well. The average cost ratio was over 1.10, with a maximum of 1.2169 and a minimum of 1.0235. In no case did it give an essentially optimal solution.

Breaking the 96 cases down into subsets based on the values of the factors, we found, as might be expected, that there was very little difference in the average or maximum cost ratios for any of the sub-groupings for either Heuristic A or Heuristic B. To a large extent, the same is true for Heuristic C, with some, but not a lot of, difference in the averages or maximums for the sub-groupings by the number of components, the C_o/C_h ratio, the C_l/C_b ratio, or the mix of desired N_i values. There were, however, much larger values for the average (1.16721), maximum (1.2169), and minimum (1.1162) ratios with P/D = 3 versus P/D = 10 (average = 1.0434, maximum = 1.0807, minimum = 1.0235) and also for the average (1.12265) and maximum (1.2169) ratios with β = 0.95 versus β = 0.75 (average = 1.08491, maximum = 1.1560).

6. Conclusion

We have proposed three heuristic solution methods as alternatives to the integerized continuous optimum solution procedure in Pentico et al. (2009b): A) use Pentico et al.'s (2009a) model to determine T and F and then use those values in the equation to determine the optimal values for the numbers of orders for each component given T and F, B) require that all components be ordered at the same times and determine the values of T, F, and N that would minimize the cost, and C) use lot-for-lot ordering of the components and determine the optimal values of T and F with that requirement. Using the same set of 96 test cases that Pentico et al. (2009b) used to evaluate the performance of the optimum-seeking model, we found that Heuristic A performed almost as well as the optimization procedure, Heuristic B performed extremely well, with an average increase over the lower bound of less than 0.1 percent, and Heuristic C, as expected, did poorly, with an average increase over the lower bound of more than 10 percent.

REFERENCES

Banerjee, A., C. Sylla, and S. Eiamkanchanalai (1990). Input/output lot sizing in single batch production systems under constant demand. *Computers in Industrial Engineering*, 19(1-4), 37-41.

Pentico, D.W., M.J. Drake, and C. Toews (2009a). The deterministic EPQ with partial backordering: A new approach. *Omega*, 37(3), 624-636.

Pentico, D.W., M.J. Drake, and C. Toews (2009b). Using the EPQ with Partial Backordering for Coordinated Planning of a Product and Its Components, *Proceedings*, National Decision Sciences Institute Conference, 2009, 2641-2646.

Sharma, S. and D.K. Singh (2004). Economic procurement cycles in response to JIT concept. *Journal of the Institution of Engineers (India)—Part PR*, 85(September), 36-40.

A copy of the full paper, including the optimization model and descriptions and performance characteristics of three heuristics, may be obtained from the second author.

LEAN SIX SIGMA FOR THE TWENTY-FIRST CENTURY

Edward D. Arnheiter
Rensselaer Polytechnic Institute, Lally School of Management & Technology
275 Windsor Street, Hartford, CT, 06120, USA
arnhee@rpi.edu
+1 860 548 7833

John Maleyeff
Rensselaer Polytechnic Institute, Lally School of Management & Technology
275 Windsor Street, Hartford, CT, 06120, USA
maleyj@rpi.edu
+1 860 548 7870

Venkat Venkateswaran
Rensselaer Polytechnic Institute, Department of Engineering & Science
275 Windsor Street, Hartford, CT, 06120, USA
venkav3@rpi.edu
+1 860 548 2458

ABSTRACT

Many organizations that have embraced the principles and techniques of Lean and/or Six Sigma are beginning to find their efforts complicated by important changes in the business landscape. These changes include full scale globalization, as well as increases in the prevalence of knowledge workers, the rate of technological change, worldwide regulatory oversight, and volatility in general. The effects of these trends on Lean Six Sigma is discussed and modifications that will need to take place in the application of Lean Six Sigma, which appear to be evolutionary rather than revolutionary, are envisioned.

INTRODUCTION

Constant, rapid change is the overarching theme for the challenges that business and management face in the new century. Business leaders have been anticipating, speaking, and writing about the demands that the new millennium would bring for several decades. While words and phrases vary, the underlying ideas are remarkably similar. There is general agreement that the winning companies will be those that convert these challenges into opportunities.

In a number of works surveying the overall change landscape, business thinkers identified challenges facing 21st century managers and organizations. The need for management to adapt to change and to find new strategies, the "New Information Revolution," the importance of knowledge worker productivity, and facing the personal challenges found in the complexities of a longer work life were identified by business guru Peter Drucker (1999) as crucial issues facing millennial business leaders. Others pointed to a volatile business environment (Hamel, 2009, Prahalad, 1998); frequently indeterminate industry and organizational boundaries (Prahalad, 1998; Hitt, 2007; Senge, 1999); global competition (Prahalad, 1998; Hitt, 2007; Nadler and Tushman, 1999); changing regulatory environments (Prahalad, 1998; Nadler and Tushman, 1999; Laurie, 2006); the opening of markets (Hitt, 2007; Nadler and Tushman, 1999); and sustainability (Prahalad, 1998). Hamel (2009) maintained that we are in an age of revolution that demands radical innovation and the harnessing of the imagination and creativity of every worker.

THE CHANGES

In this section, the main trends expected to affect the practice of business management are categorized. The authors wish to thank Linda Knaack for her helpful assistance in the development of this section.

Sustainable Business Development

An organization's ability to be sustainable will be a key factor in determining its success in the new century (Rainey, 2006). Industrial growth is facing increasing social and environmental limitations. Peter Senge is one of the major proponents of a new environmentalism driven by innovation rather than regulation and views sustainability as an opportunity for businesses to develop new technologies, products, processes, and business models as they transition from mere compliance to full sustainability (Senge et al, 2001; Senge, 2006; Senge, 2008; Nidomolu et al, 2009). In high demand will be leaders who can effectively spearhead efforts to create firms that are sustainable (Rainey, 2006; O'Toole and Bennis, 2009).

The Global Marketplace

In a volatile world, successful companies will pay attention to agility while continuing to focus on quality, cost, and efficiency. Doz, Santos, Williamson, & Berrebi (2004) argued that converting Europe's relatively small geographic area and its diversity from obstacles to advantages demands new ways of thinking about how companies innovate. Their meta-national innovation processes (sensing, mobilizing, leveraging) should be applicable to countries around the world.

Global Operations

A highly volatile nation, India is a good model for operating in an environment of "shifting regulations, spurts of growth, capital shortages, and challenging supply base" (Prahalad, 2009). Many corporations are realizing that the knowledge they need to innovate is often found far afield, and globalizing innovation is becoming a cost effective option. Companies like Airbus and STMicroelectronics have demonstrated that increased innovation at lower costs gives them a competitive edge. Globalization demands an entirely new mind-set, not simply a change in processes (Santos et al, 2004).

Rapidly Evolving Technologies

Thomas W. Malone, the Patrick J. McGovern Professor of Information Systems at MIT's Sloan School (2009), when he announced the research and educational initiative *Inventing the*

Organizations of the 21st Century maintained that "We are now at the threshold of a new era, driven this time not by the technologies of production and transportation, but by the technologies of information, communication, and coordination. These technologies…hold the potential to completely transform the nature of work throughout the world." Specific technological advances like e-commerce, virtual teams, and artificial intelligence will make the manager's job more complex and consequently more difficult. The increasing rate of this technological change will result in smaller product life cycles with little time to recoup research and development costs. Also, while technology facilitates gathering knowledge, it also greatly increases the volume acquired. The development and acquisition of knowledge is an important piece of competitive advantage (Hitt, 2007; Hill, 2004).

Changing Regulatory Environments

Both deregulation, which lowers geographical and other entry barriers to many industries, and increased regulations, in the area of sustainability, exert significant pressures on any firm operating in a global economy (Prahalad, 1998,Senge, 2008; Nidomulo et al, 2009).

Rise of the Knowledge Worker

Knowledge workers use information or develop knowledge in the course of their jobs. They represent an ever increasing segment of the work force, and managing them effectively is crucial to any organization wishing to compete successfully in the new millennium. The knowledge worker exhibits an attitude towards work that is different from traditional workers and companies wishing to increase productivity must motivate rather than control them (Drucker, 1999; Kanter, 2009).

Managing in a Volatile Environment

"Over the years managers have developed tools and techniques to overcome challenges ranging from inconsistent quality to stagnant productivity (e.g., Six Sigma, Total Quality Management, and just-in-time supply chains). Now what they need is a system for addressing volatility. Prahalad (2009) asks "How does a chemical company, for example, cope with oil prices that bounce from $50 a barrel to $150 and back in 18 months?" This is a challenge to organizations of the new century that many business thinkers and business leaders are working hard to address (Hamel, 2009; Prahalad, 1998; Prahalad; 2009; Hamel, 2000).

EFFECTS ON LEAN SIX SIGMA

The American Society for Quality (ASQ) Six Sigma body of knowledge was used as a basis for documenting the effects that the changes categorized above will have on the implementation of a process improvement program, such as Lean Six Sigma. This structure is justified because the ASQ body of knowledge contains Lean methods and principles, and because many organizations that operate successful process improvement programs use a variety of names for the program and emphasize different aspects of Lean and/or Six Sigma (Maleyeff, 2007).

The ASQ body of knowledge includes a number of main sections. Using a series of tables, details are provided regarding: (a) how the current elements of Lean Six Sigma mesh with the upcoming changes, and (b) what needs to change in the application of Lean Six Sigma in order to remain relevant. Tables 1-5 (respectively) concern Enterprise-Wide Deployment; Organizational Process, Management, and Measures; Team Management; Define-Measure-Analyze-Improve-Control (i.e., DMAIC, which is the main project organization model used within Six Sigma); and Design for Six Sigma.

Table 1
Enterprise-Wide Deployment

	How Lean Six Sigma Elements Mesh	What Needs to Change in Lean Six Sigma
Sustainable Business Development	Change management skills and knowledge of Lean Six Sigma principles will continue to be crucial in meeting the goal of sustainability.	A better and systematic way to define the value of long term sustainability will be necessary, including cap and trade issues.
The Global Marketplace	The strong focus on understanding customer needs will be important as the diversity in customer types increases.	Defining value in various economic and cultural systems will complicate project selection, including effects of tariffs.
Global Operations	Leadership and knowledge in the areas of Lean and Kaizen will remain important skills for global operations managers.	Motivation for change will vary across economic and cultural systems, and this will complicate implementation of new processes.
Rapidly Evolving Technologies	The ability to understand how to integrate technology within business processes will take on increased importance.	The integration of technology across physical and cultural boundaries will be critical, and success may depend on learning from others.
Changing Regulatory Environment	Change management skills and solid leadership are critical when the regulatory environment is volatile.	Improvement project implementation may be complicated by limitations based on governmental laws and regulations.
Rise of Knowledge Workers	The ability to understand why and how Lean Six Sigma works will facilitate better understanding of its implementation.	Improvement efforts may move from process improvement to creation of effective working environments, including the virtual workplace.
Managing in a Volatile Environment	Change management skills are critical when business conditions and supply chains are unstable and volatile.	Project durations will need to be short but thorough, with many interconnected projects undertaken simultaneously.

Table 2
Organizational Process, Management, and Measures

	How Lean Six Sigma Elements Mesh	What Needs to Change in Lean Six Sigma
Sustainable Business Development	Although measurements are important, many traditional business performance and financial measures do not address sustainability.	Financial measures will need to change to create an emphasis on how improvements impact societal and environmental goals.
The Global Marketplace	Emphasis on critical-to-quality and related measurements will assist with the integration of performance metrics.	Decisions must take into account different market perspectives and how services are managed in diverse cultures.
Global Operations	Lean Six Sigma is applied and will impact all operations, including domestic and overseas facilities.	Terminology of performance measures may not be similar and how they motivate action across counties will vary.

Rapidly Evolving Technologies	Rapidly evolving technologies might require new critical-to-quality and other performance measures.	Financial measures need to focus more directly on the impact of technology and management must adapt fast to changes.
Changing Regulatory Environment	Current efforts to implement Lean Six Sigma include critical stakeholders, so adding another stakeholder should be easy.	More key stakeholders (government, regulators) will exist and intellectual property issues will increase in complexity.
Rise of Knowledge Workers	Business measures and metrics are important tools for knowledge workers, but emphasis must be placed on worker satisfaction.	Improvements are needed in how performance of knowledge workers is measured and used to develop allegiance.
Managing in a Volatile Environment	Balancing the need for longer-term financial measures with a rapidly changing business environment will be challenging.	Cause-and-effect will be more difficult to determine and certain risks (political, legal, etc.) will increase.

Table 3
Team Management

	How Lean Six Sigma Elements Mesh	**What Needs to Change in Lean Six Sigma**
Sustainable Business Development	Teams might require retraining to focus their efforts on long-term issues and projects that contribute to sustainability.	Requires involvement of a green-certified member and more stakeholders (e.g., energy providers, community, or government).
The Global Marketplace	Team management approaches using diverse, international team members can enable a company to stay in tune with market.	Need to include members on teams that understand cross-region markets and mores, perhaps using localized sub-teams.
Global Operations	Team skill sets required and developed in black belt system would allow people to work well on “virtual teams” across the globe.	Challenges need to be overcome with multi-functional teams in disperse locations, having different languages and cultures.
Rapidly Evolving Technologies	Traditional team process will likely work well, because it capitalizes on the knowledge power of the group versus individual.	Need to expand team member scope with those that understand potential and limitations of technology solutions.
Changing Regulatory Environment	Teams must be agile and able to quickly communicate any changes in foreign and domestic government regulations.	New team members from regulatory bodies will be needed with team-wide awareness of regulatory effects across the globe.
Rise of Knowledge Workers	Concept of rewarding teams rather than individuals will become more difficult as good workers become harder to retain.	Need to convince workers that real benefits will come to them personally since they will not be motivated by slogans or money.
Managing in a Volatile Environment	Traditional team management approaches might be too slow and cumbersome, lacking the agility necessary to adapt.	Many members of the team will be new to the company or to the job and therefore teams need to be amorphous.

Table 4
The DMAIC Process

	How Lean Six Sigma Elements Mesh	**What Needs to Change in Lean Six Sigma**
Sustainable Business Development	DMAIC can be applied to any process – therefore it could be used as part of a broad intangible framework like sustainability.	Need to find better ways to measure future and indirect effects, and would need to sacrifice short-term profits for sustainability.
The Global Marketplace	This logical and thorough “scientific method” can be well suited for competing within the global marketplace.	Need to be proactive in identifying opportunities for improvement as market preferences change and evolve.

Global Operations	DMAIC is useful for all operations, whether domestic or overseas; the universality of the mathematical approaches is appealing.	Differences in education practices could complicate training of the tools; challenges in central vs. local control need to be overcome.
Rapidly Evolving Technologies	DMAIC may be too slow and ponderous when technology is moving at breakneck speed with projects taking months to complete.	More risk taking in choosing projects and potential for "user" of project results being unable to understand their implementation.
Changing Regulatory Environment	The DMAIC process can provide a structured way for companies to assess changes within the regulatory environment.	Changes in regulations will often motivate improvement projects but recommendations may require law changes to implement.
Rise of Knowledge Workers	DMAIC helps workers gain in-depth insight into the details of the process – not just superficial "managerial" knowledge.	Tools will evolve to downplay manufacturing by including those specific to knowledge jobs, and this would impact training regimens.
Managing in a Volatile Environment	DMAIC may be too slow and ponderous when technology is moving at breakneck speed with projects taking months to complete.	Need to modify DMAIC to move quickly (Kaizen approach would help) and controls will be emphasized to identify new problems.

Table 5
Design for Six Sigma

	How Lean Six Sigma Elements Mesh	What Needs to Change in Lean Six Sigma
Sustainable Business Development	The tools and concepts of DFSS will take on increased importance, as the criteria for effective designs are broadened.	Norms, guidelines, and paradigms for sustainable design need to be part of DFSS projects to account for total life cycle costs.
The Global Marketplace	The tools and concepts of DFSS will take on increased importance, as the diversity of customer types increases.	Designs must be cognizant of various user emphases with decisions regarding design flexibility becoming important.
Global Operations	The tools and concepts of DFSS will take on increased importance, as the diversity of operations types and locations increases.	Designs must take into account manufacturability in more diverse ways, including capabilities across regions.
Rapidly Evolving Technologies	DFSS would still be valid, but the methodology needs to be applied in a flexible, rapid manner.	Project teams must be aware of how users from around the globe understand or deal with technology.
Changing Regulatory Environment	The tools and concepts of DFSS will take on increased importance, as the number of stakeholders increases.	Designs must be developed that are able to evolve as regulations change and are updated.
Rise of Knowledge Workers	The black belt system requires high skill set; black belts would do well in a knowledge worker economy.	An increased emphasis will be placed on the design of services and solutions, with corresponding differences in approaches.
Managing in a Volatile Environment	DFSS would still be valid, but the methodology needs to be applied in a flexible, rapid manner.	Design processes must be fast and their work processes will need to be periodically reviewed and updated.

CONCLUSIONS

Given the broad nature of this work and the potential for a widely diverse set of opinions, the authors look forward to a spirited debate on the future prospects of Lean Six Sigma.

REFERENCES

References available upon request from John Maleyeff (contact information listed above).

THE EFFECT OF TECHNOLOGY AND QUALITY ON HOSPITAL FINANCIAL PERFORMANCE: ASSESSMENT OF HOSPITAL DIRECTOR OPINIONS IN ANKARA PROVINCIAL CENTRE

Gamze Y. TARCAN: MSc., Hacettepe University, Health Services Technical School, Ankara, Turkey. E-mail:gamze@hacettepe.edu.tr, Phone: +90 538 825 88 15.

Menderes TARCAN: Ph.D., Ministry of Health, Performance Management and Quality Improvement Directorate, Ankara, Turkey. E-mail:mtarcan@hacettepe.edu.tr, Phone: +90 533 307 51 32.

Hacer OZGEN: Ph.D., Hacettepe University, Faculty of Finance and Administrative Sciences, Health Administration Department, Ankara, Turkey. E-mail:hozgen@hacettepe.edu.tr, Phone: +90 538 561 78 12.

ABSTRACT

The aim of this study is to determine the effect of technology and quality on hospital financial performance.

The factors that affect hospital financial performance are found below: the most important variable is clinical technology. In respect of this; investments for clinical technology is the most important factor that affects hospital financial performance positively. The other factor that affects the hospital financial performance is process quality. Another meaningful variable affects hospital financial performance is ownership and it can be said if that the ownership is SSK compare to university hospitals that are reference affect more positively hospital financial performance.

Key words: Technology, quality, hospital financial performance.

Introduction

Last two decades more hospitals developed formal technology assessment processes in an attempt to control costs while improving both clinical and process quality. But there is limited research and practical guidelines on how hospital technology, quality, and financial performance are related. The aim of this research is to determine the effect of technology and quality on hospital financial performance in the health sector of Ankara province in Turkey.

Material and Method

The content of this survey contains Ministry of Health, Social Insurance Organization, University and private hospitals' director, vice director, hospital manager, vice manager, revolving fund accountants and head nurse and vice head nurse which is in Ankara Provincial Centre. Because hospital investments, products and services are purchased through the revolving fund revenue, accountants are included in order to provide data about hospital financial performance. No sample choosen for the research, to reach the total field under the survey is aimed.Total field under survey is 510 directors, 141 directors replied the overall survey, as a result can be worked with %27.64 of the total field.

The survey was used to collect data in this research was evolved by Li and Collier [3], has been translated into Turkish and adapted descriptive questions. Survey is composed of two parts. There are descriptive questions about directors in the first part. In the second part, there are questions about technology, quality and hospital financial performance. One of every category includes related questions. In the survey form; 1; “never investment”, “never importance”, “being very low” symbolized, and 7; “great investment”, “great importance”, “being very high” symbolized.

The directors’ demographic characteristics have been analized by descriptive statistical methods like as frequency and percentege distribution. In this research factor analysis which is a multivariate analysis method, was used to find directors’ assessments about technolgy, quality and financial performance. The factor analysis was used determining the dimensions of the scale. And multi regression models were used to determine the predictors of the financial performance and the other dependent variables.

Findings

Table 1 shows some descriptive statistics of the hospitals included in this study. 42.55 % (n=60) of total hospitals directors (who have been applied for their opinion) are female, 57.45 % (n=81) are male. The distribution of hospitals by the ownership is; 26.24 % (n=37) Ministry of Health hospital, 27.66% (n=39) university hospital, 24.82% (n=35) Social Insurance Organization hospital and 21.28% is private hospital.

Table 1:Descriptive Statistics of Sample

	N	%
Gender		
Female	60	42.55
Male	81	57.45
Ownership Of Hospitals		
Ministry of Health	37	26.24
University	39	27.66
Social Insurance Organization	35	24.82
Private	30	21.28
Functions Of Directors		
Medical Director	6	4.26
Vice Director	24	17.02
Hospital Manager	12	8.51
Vice Manager	63	44.68
Revolving Fund Accountant	6	4.26
Head Nurse	12	8.51
Vice Head Nurse	18	12.77
TOTAL	**141**	**100**

Table 2 shows the effects of ownership and clinical quality on the clinical technology. The Clinical technology variable affects clinical quality statistically significant (β=0.428; =5.393 and p<0.01).

Table 3 shows the results of regression analysis about factors that affect process quality . There is a positive relation between information technology and process quality (β=0.448;

t=5.934 and p<0.01). Other significant variable which affects process quality in the model is ownership of the hospital. If the ownership is Social Insurance Organization, the process quality has been affected negatively as to reference hospitals (β= -0.183, t=-2.067 and p<0.05); if the ownership of the hospital is private, it affects process quality positively as to reference group hospitals (β=0.168; t=1.879 and p<0.1).

Table 2. The Effect of Ownership and Clinical Technology on Clinical Quality, Summary of Regression Results

	β	t	p	VIF
Clinical Technology	0.428	5.393	0.000*	1.092
Ministry Of Health	0.079	0.857	0.393	1.457
Social Insurance Org.	-0.066	-0.723	0.471	1.428
Private	0.086	0.947	0.345	1.443
(Ref: University)				
R^2	0.214			
F	9.276			
Sig.	0.000			

Dependent Variable: Clinic Qualitiy, *= p=<0.01; ** p=<0.05; *** p=<0.1

Table 3. The Effect of Information Technology and Ownership on Process Quality; Summary of Regression Results

	β	t	p	VIF
Information Technology	0.448	5.934	0.000*	1.116
University	0.028	0.316	0.752	1.500
Social Insurance Org.	-0.183	-2.067	0.041**	1.525
Private	0.168	1.879	0.062***	1.572
(Ref:Ministry of Health)				
R^2	0.304			
F	14.869			
Sig.	0.000			

Dependent Variable: Process Qualitiy, *= p=<0.01; ** p=<0.05; *** p=<0.1

In the last stage, all the variables; clinical quality, process quality, clinical technology, information technology and ownership are analized at the same time and tried to examine that how has these variables affected financial performance.

Table 4. Factors Effect Hospital Financial Performance; Summary of Regression Results

	β	t	p	VIF
Clinic Technology	0,373	2.682	0.008*	3.017
Information Technology	-0.114	-0.78	0.437	3.348
Clinic Quality	-0.095	-0.92	0.360	1.659
Process Quality	0.196	1.714	0.089***	2.031
Ministry of Health	-0.061	-0.63	0.530	1.48
Social Insurance Org.	0.179	1.767	0.080***	1.598
Private	0.043	0.44	0.661	1.506
(Ref:University)				
R^2	**0.145**			
F	3.234			
Sig.	0.003			

Dependent Variable: Hospital Financial Performance, *= p=<0.01; ** p=<0.05; *** p=<0.1

Table of 4. shows the results of the regression analysis. As seen from the table; Process quality affects hospital financial performance and it is statistically significant ($\beta=0.196$; $t=1.714$; $p<0.1$). Clinical technology was the important variable that affects hospital financial performance ($\beta=0.373$; $t=2.682$; $p<0.01$). Other significant variable is ownership; If the hospital ownership is Social Insurance Organization, it affects hospital financial performance positively as to reference group of university hospitals ($\beta=0.179$; $t=1.767$; $p<0.1$).

Conclusions and Recommendations

The results have been found in the research which can be tried to determine the effect of technology and quality on hospital financial performance;

- Sample survey which was evolved by Li and Collier [3], has been translated into Turkish was used to collect data in this research. It's reliability has been tested and it was found reliable. At the end of factor analyze it was determined that survey has been classified into five dimensions. These are; cilinical quality, process quality, clinical technology, information technology and hospital financial performance (Appendix).
- We saw that the most important variable which affects clinical quality in the hospital is clinical technology.
- The factors that affect process quality are; investment on information technology and ownership of the hospital. Investment on information technology affects process quality positively. And If the ownership is Social Insurance Organization, it affects hospital process quality negatively as to reference group of Ministry of Health hospitals. If the ownership is private, it affects hospital process quality positively as to reference group of Ministry of Health hospitals.
- The factors which affects hospital financial performance are;
- The most important variable that affects hospital financial performance is clinical technology. According to this, investments on clinical technology affect hospital financial performance positively.
- The other factor that affects hospital financial performance significantly is process quality.
- Another significant variable is ownership that affects financial performance of the hospitals. If the hospital's ownership is Social Insurance Organization , it affects hospital financial performance positively as to reference group of university hospitals.

In light of the results of this study, it is possible to claim that;

A good financial performance is a prerequisite for the survival of the hospitals, therefore it can be usefull that the hospitals examine the factors which affects their hospital financial performance. Following the development of hospital clinical technologies, the hospitals should renew their clinical technology certain extent, may be useful in terms of financial performance improvements. Technological investments constitutes a large part of hospital costs [2]. To find out the effects of investment and ascending quality on the financial performance can be guide for the hospital management strategy. In order to survive in the competitive markets ,hospitals should increase the level service and process quality, invest on optimal technology. Investing the information systems in the hospitals provide the efficienct cost accounting system for the hospital [1] and limited resources of the hospitals can be used efficiently. In conclusion, to having poor financial performance, means limited resources are used non-productively. The hospitals must develop administrative skills which improve their financial performance, and should emphasize all factors which affect their financial performance.

APPENDIX: STUDY SCALE ITEMS, FACTOR LOADINGS AND RELIABILITY

Study Scale Items	Factor Loadings				
	Process Quality	Information Technology	Clinical Quality	Clinical Technology	Hospital Financial Performance
Laboratory equipment and information systems				0.79	
Radiology equipment and information systems				0.71	
Pharmacy information systems				0.71	
Automated drug dispensing system				0.52	
Digital compressed video technology		0.81			
Teleradiography		0.78			
E-mail system		0.72			
Paging system		0.67			
Computers in doctor's office		0.62			
Fax machine in every department		0.62			
Improve clinical outcome			0.72		
Customer satisfaction about clinical outcome			0.54		
Department heads coordinate inter-departmental and intradepartmental information resource decisions and activities	0.83				
Giving staff members more quality control responsibility	0.83				
Assigning a borader range of tasks to staff members	0.82				
Implementing employee empowerment programs	0.82				
Department heads integrate the departments' information resource plan with the organization's plan	0.76				
Providing supervisor training for improving leadership capabilities	0.76				
Creating oppurtunities for staff members to learn and use skills that go beyond current job assignment	0.73				
Staff members effectively use all equipment in their activities	0.71				
Increasing the level of job enlargement	0.71				
Staff members have the knowledge and skills required to perform their jobs	0.7				
Improve productivity of medical&nursing staff	0.69				
Provide training to improve staff member's clinical competence	0.69				
Establish partnerships with other hospitals and health care networks to share training and spread job opporunities	0.67				
Top management assures that organization staffs are trained in managing information resources and member's clinical competence	0.57				
Attract physician to your hospital	0.56				
Forge partnership with physician	0.51				
Top management fosters communication among different departments to coordinate information resource management activities	0.48				
Return on investment					0.91
Operating profit					0.88

Factors	Cronbach α
Clinical Technology	0.84
Information Technology	0.87
Clinical Quality	0.82
Process Quality	0.93
Hospital Financial Performance	0.85
General Reliability of the Scale	**0.93**

References

[1] Kısaer, H. "Activitiy based costing in the hospitals." *Journal of Productivity,* 1996, 3 (1), 151-160.
[2] Yeğinboy, Y. *Assessment of Financial Performance in the hospitals.* Dokuz Eylül Üniversitesi, 1992.
[3]. Li, L.X., Collier, D.A. "The Role of technology and quality on hospital financial performance: An Exploratory analysis." *International Journal of Service Industry Management*, 2000, **11,** 202-224.

THE IMPACT OF CULTURE ON CESAREAN RATES:
A CROSS-COUNTRY STUDY

Vasanthakumar N. Bhat, Lubin School of Business, Pace University (212) 346-1200, vbhat@pace.edu

ABSTRACT

The values and beliefs of a country do have impact on the medical practice. In this paper, we examine the relationship between cesarean rates and socioeconomic, demographic, healthcare, and cultural variables. We use the Spearman Rank correlation and multiple regression analysis to study the variations in cesarean rates in 25 countries belonging to the Organisation for Economic Cooperation and Development (OECD). Our study indicates a statistically significant association between cesarean rates and masculinity and uncertainty avoidance indicators, developed by Hofstede, of a country. The influence of culture on medical decisions calls for training physicians and healthcare professionals about the impact of beliefs and values on their medical practice.

INTRODUCTION

Pregnancy and birth-related medical expenditures account for about seven percent of total medical expenditures in the United States in 1993 (HIAA, 1995). Pregnancy and birth-related reasons account for the most frequent cause of hospital admissions, constituting about four million births annually. Therefore, birth-related expenditures are of significant concern to health administrators and policy makers. High rates of cesarean sections are of concern to those who want to reduce health care costs. Long recovery times, high morbidity rates, and increased medical and surgical costs have caused health care professionals to critically evaluate cesarean section rates. Berwick(1994), for example, includes reduction of cesarean rates to pre-1980 levels among his eleven worthy goals for clinical leadership of health system reform in the United States.

We examine the influence of nonclinical factors on cesarean rates in this paper.. Our focus of analysis is on cesarean rates in countries belonging to the Organisation for Economic Co-operation and Development (OECD). In addition to examining the impact of demographic, economic, and health care factors, we also analyze the influence of culture on cesarean deliveries. Even though we may not provide a definite answer to the relevance of culture to cesarean deliveries, our analysis indicates that cultural attitudes may explain a portion of variations in cesarean section rates among different countries.

This paper is organized into 5 sections. In the next section, we examine trends in cesarean deliveries over time and among different countries. We describe reasons for considering culture in explaining variations in cesarean section rates. We then present univariate associations between cesarean rates of various countries and socioeconomic, demographic, and cultural factors. We also present a multivariate analysis of cesarean rates. We close this paper with conclusions.

INTERNATIONAL TRENDS IN CESAREAN RATES

There are not many studies that examine variations in cesarean rates across countries. Belizan et al. (1999), after comparing cesarean rates in 19 Latin American countries, conclude that cesarean rates are positively associated with gross domestic product per capita and that private hospitals have higher cesarean rates than public ones.

Cesarean section rates vary significantly over time within and among the countries we examine in this paper. Cesarean rates from 1970 for 17 countries belonging to OECD are given in Figure 1. The graphs indicate that trends in cesarean rates are up for most countries from 1970. The cesarean rates for most countries in recent years have fallen between those of the Netherlands and Italy. For the Netherlands, the cesarean rates varied between 65 per 1000 live births in 1985 to 129 per 1000 live births in 2000. For Italy, cesarean rates varied between 112 per 1000 live births in 1980 to 347.6 per 1000 live births in 2001. For the United States, the cesarean rates were almost stable between 207 per 1000 live births in 1996 to 228 per 1000 live births in 1999. The cesarean rates were only 28 per 1000 live births in 1960 in the United Kingdom. However, the rates rose to 124 in 1990 and 170 per 1000 live births in 1997. The cesarean rates for various countries for around 1999 are given in Figure 2. The cesarean rates vary significantly with 113.4 per 1000 live births in the Netherlands and 323.9 per 1000 live births in Italy.

Culture and Medical Practice

There are many definitions for culture. Hofstede (1994, p.260) defines culture as "the collective programming of the mind which distinguishes the members of one group or category of people from another." Religious beliefs, food preferences, aesthetic choices, and attitude towards authority are some examples of mental programming. Symbols represent the most specific and values represent the most general manifestation of culture. Culture is learned rather than inherited. Culture helps to differentiate a member of one group from another. On a basic level, culture represents the mode of operation of a group and the value system that underlies it. Cultural differences reflect different values. Values represent preferences for certain outcomes over others. The differences between cultures do not mean that every one within one culture shares the same mental programming. National cultures represent majority mental programs shared by the middle classes in each country. Hall (1984) presents a different perspective on culture. According to Hall, culture acts as an invisible mechanism that controls our thoughts. We internalize the cultural norms and behave within their boundaries. Even though culture is arbitrary, we feel that they are real. Spencer-Oatey (2000) includes attitudes, beliefs, behavioural norms, assumptions, and values within the definition of culture. These are shared by group members and used by them to influence the behaviour of other members of the group and to interpret the meaning of other people's behaviour. To sum up, culture can be thought of as the attributes shared by a group that form the basis for behavior.

Even though a number of authors have put forth a universally applicable basis for classifying cultural patterns, the framework provided by Dutch organizational anthropologist Hofstede(1980) stands out. Even though his framework is very general, it is easy to apply to a variety of situations. His framework also helps to represent a culture using four dimensions. Hofstede came up with four dimensions of culture by examining work-related values in employees of IBM during the seventies. According to Hofstede (1994), the cultures of a country can be characterized by four independent dimensions: power distance, individualism /collectivism, masculinity/femininity, and uncertainty avoidance.

Power distance provides a measure of societal inequality. It quantifies the degree to which less powerful members expect and accept that power is distributed unequally. (Hofstede, 1994, p. 28). Austria, Denmark, and New Zealand have the lowest and France, Turkey, and Belgium have the highest values for power distance among the countries we examine in this paper.

The second dimension is about alone versus together. Individualism is related to societies in which everyone is expected to care for himself and his immediate family. Collectivism, on the other hand, pertains to societies in which people from birth belong to cohesive groups and members are protected for their loyalty (Hofstede 1994, p. 51). This is one of the most researched dimensions of culture. Portugal, Greece and Turkey have the lowest values and USA, Australia, and the UK have the highest values for individuality.

The masculinity/femininity dimension deals with toughness and tenderness. In a masculine society, men are assertive, tough, and concentrate on material success whereas women are modest, tender, and deal with the quality of life. In a feminine society, gender roles are not that distinctive. In a feminine society, both men and women are modest, tender, and concerned with the quality of life (Hofstede, 1994, p. 82-3). Denmark, Sweden, and Ireland have the highest values and Belgium, Portugal, and Greece have the lowest values for this dimension.

Uncertainty avoidance represents the degree to which the members of a society are threatened by uncertain situations (Hofstede, 1994, p. 113). This indicator is found to be positively associated with anxiety levels in various countries. Sweden, Norway, and the Netherlands have the lowest and Italy, Australia, and Japan have the highest values for uncertainty avoidance.

According to McPherson (1990), medical practice patterns vary significantly across and within countries even after taking into account variables such as age, morbidity, and institutional arrangements of the health delivery system. Patient needs alone cannot explain differences in procedures and their utilization rates among regions, hospitals and physicians in the United States. (Wennberg & Cooper, 1999). The average number of medicines prescribed varies significantly across countries (OECD, 2003). Therefore, factors other than patient needs play a role in medical practice.

The impact of culture on medical practice has not been extensively examined. Draguns and Tanaka-Matsumi (2003), after reviewing literature on the impact of culture on psychopathology, suggest further research to evaluate the association between psychological disturbances and culture. They also propose hypothesis relating to the cultural dimensions of Hofstede (1980) and their possible relationship to psychiatric symptomatology. According to Arrindell et al.(2003), Hofstede's national masculinity index is related to national depression levels based on an analysis of data collected in 14 nations in Europe.

Rising cesarean rates have caused researchers to study non-medical factors that contribute to the use of medical interventions during birth. A variety of physician-related attributes such as medical training, experience, working schedules, and medical practice organizations have been examined to identify factors that influence cesarean section rates. (Barros et al.,1996) (Burns et.al, 1995). Some researchers have examined why women passively accept cesarean section when it is clearly known that cesarean sections are high-risk. These researchers (Jordan, 1978), (Davis-Floyd, 1990, 1993), (LoCicero, 1993) argue that belief among women that technology is superior to nature and masculine dominance prevents women from objecting to cesarean procedures.

METHOD

The purpose of this paper is to examine empirically whether there is any association between cesarean rates and culture. The cesarean rates are from the database published by Organisation for Economic Cooperation and Development (OECD, 2003). This interactive health database contains data on a variety of major indicators of health care systems of 30 OECD member countries. Data on surgical procedures are presented according to the classification ICD-9-CM. The cesarean rates are for ICD codes 74.0-74.2, 74.4, 74.99. The cesarean rates are calculated for 1000 live births. We represent culture using the four dimensions of culture proposed by Hofestede(1980). Even though we examine 25 countries in this paper, the multiple regression analysis is based on the following 16 countries:

Australia, Finland, the Netherlands, United States. Austria, France, Norway, Belgium, Germany, Sweden, Canada, Ireland, Turkey, Denmark, Italy. the United Kingdom

The data for cesarean rates are from OECD Health Data 2003 3rd Edition. The cultural dimensions are from Hofestede (1991).

We first perform univariate analysis between various socio-economic factors and cesarean rates. For this purpose, we use the Spearman Correlation because it is not affected by extreme values and does not require any distributional assumption. However, the results from the Spearman Correlation should be viewed with caution, as it does not take into account the effect of other variables. We therefore also present multivariate analysis using multiple linear regression analysis between cesarean rates and cultural dimensions.

RESULTS

Univariate Analysis

We perform the Spearman Rank Correlation between cesarean section rates and various demographic, socio-economic, and cultural factors. The results are presented in Table 1. We calculate the Spearman rank correlation between cesarean rates and variables relating to mother's and infant's health indicators.

Multivariate Analysis

Since a univariate analysis considers the effect of only one variable at a time, we examine the impact of several variables on cesarean rates using the multiple linear regression analysis. We use cesarean rates as a dependent variable and per capita GDP, power index, individuality index, uncertainty avoidance index, masculinity index, physicians per 1000 people, and income disparity (Gini coefficient) as independent variables. We start the estimation process with all seven variables and then reduce successively by removing not significantly correlated (on the basis of individual t-statistics) with the dependent variable. The regression results are presented in Table 2.

We find that our regression results explain more than 75 percent of variations in cesarean rates. We find that GDP per capita, power index, and individuality index are not statistically significant at the 5-percent level. The restricted model without these variables indicates that physicians per 1000 people, income disparity expressed as a Gini coefficient, and masculinity index are positively related and uncertainty avoidance index is negatively related to cesarean rates.

DISCUSSIONS

Cesarean rates are considered significantly higher than medically necessary. According to Shearer (1993), medically necessary rate is about 12 percent. Except the Netherlands, all countries we examined in this paper have cesarean rates higher than 12 percent. Many mothers and babies are therefore subjected to unnecessary risk of medical, behavioral, and psychological complications. In spite of serious efforts by various medical organizations and groups, cesarean rates have not been substantially reduced. Therefore, this problem needs to be examined from a variety of perspectives. We provide a new perspective in this paper examining the impact of culture on the variations in cesarean rates across countries.

Our analysis shows that cesarean rates are positively associated with per capita physicians. This is consistent with the conclusion reached by Belizan et al. (1999), using univariate analysis for Latin American countries. The Gini coefficient represents income disparity. High family income and access to private health insurance are major parameters of use of technology during birthing. Since a higher Gini coefficient represents higher income inequalities, countries with higher Gini coefficients are likely to have higher cesarean rates because women with high income can afford cesarean deliveries.

Our analysis shows that masculinity index is positively related to cesarean rates. Countries with high masculinity index are likely to value assertiveness, achievement, and manliness. Low masculinity countries tend to focus on quality of life, sympathy for the unfortunate, and helping others. Low masculinity countries prefer equality between males and females. Sex roles in high masculinity countries are distinct and gender inequality is considered beneficial. Therefore, a positive relationship between masculinity index and cesarean rates provides empirical basis for the assertion by LoCicero (1993) that

Table 1: The Spearman Rank Correlation between Cesarean Rates and Various Demographic, Economic, and Cultural Factors

Variable	Spearman Rank Correlation	Level of Significance	Number of observations
Infant mortality Death/1000 live births	+0.2078	0.32	25
Maternal mortality Death per 100,000 live births	-0.3086	0.13	25
Low birthweight % of total live births	+0.3093	0.13	25
Obese population % females, BMI>30kg/m2	+0.3401	0.11	23
Female practicing physicians %	-0.2079	0.33	24
% females in lab. Force	-0.2666	0.20	25
Per Capita GDP US$	+0.0562	0.79	25
Per Capita GDP US$ PPP	+0.1577	0.45	25
Gini coefficient	+0.645	0.005	17
Physicians per 1000 people	-0.089	0.67	25
Power Distance Index	+0.215	0.38	19
Individualism Index	+0.153	0.53	19
Uncertainty Avoidance Index	+0.150	0.54	19
Masculinity Index	+0.6178	0.0048	19

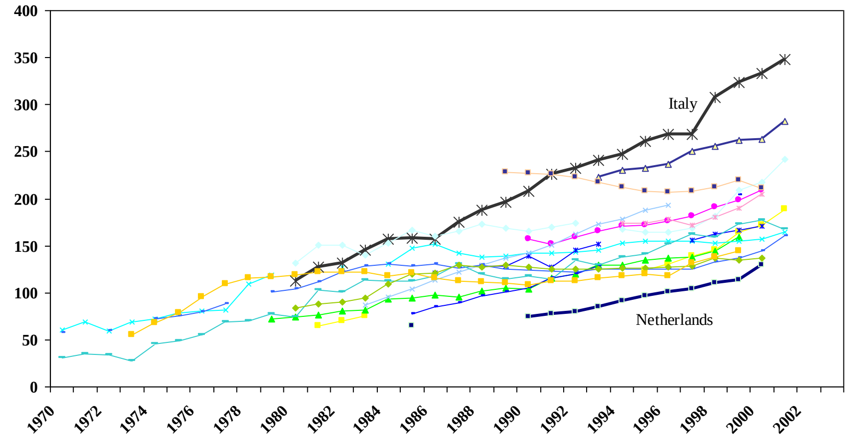
Table 2: Multiple Linear Regression Analysis with Cesarean Rates as a Dependent Variable

Independent Variable	Coefficient	t-value	Probability Level	Coefficient	t-value	Probability Level
Intercept	-120.872	-0.866	0.412	-130.439	-1.722	0.113
Per Capita GDP (US$)	-0.002	-0.666	0.524			
Power	0.048	0.036	0.972			
Individuality	0.698	0.598	0.566			
Uncertainty Avoidance	-1.119	-1.051	0.324	-1.1378	-2.476	0.031
Masculinity	1.063	1.751	0.118	1.1396	3.201	0.008
Physicians per 1000 people	57.894	3.117	0.014	58.1373	3.773	0.003
Income Disparity (Gini coefficient)	4.490	1.6	0.148	5.2366	2.773	0.018
R-squared	0.780			0.75		
R-squared (adjusted)	0.585			0.659		
F	4.02			8.25		
Probability	0.035			0.0025		
Number of observations	16			16		

male dominance is a cause for excessive cesarean rates in the United States. LoCicero claims that the male dominated obstetric system does not serve the needs of the women. The training and socialization of obstetricians continue to instill male dominance, and therefore requires radical changes to reduce cesarean rates.

The uncertainty avoidance index affects cesarean rates negatively. The uncertainty avoidance index is a proxy for the anxiety level of a culture. This index represents the degree to which people are anxious about unknown situations. However, anxiety is not the same as fear. Unlike fear, which has an object, anxiety does not have any object. In high anxiety avoidance countries, people are much more anxious than in low anxiety avoidance countries. According to Hofstede(1991, p.116), high uncertainty

Figure 1: Trends in Cesarean Rates per 1000 Live Births, 1970-2002



avoidance countries abhor confusing situations. They prefer structures that make events predictable. As a result, high uncertainty avoidance countries are willing to engage in risky behavior to reduce ambiguities. Based on this reasoning, the uncertainty index should be positively related to cesarean rates, contradicting our conclusion from regression analysis. However, there are conflicting findings about the impact of anxiety on cesarean procedures. According to Ryding et.al.(1998), based on a study in Sweden, fear of childbirth during the third trimester of pregnancy increases likelihood of emergency cesarean section. However, according to Slade and Johnson(2002), a study in the United Kingdom, reached a different conclusion. Sweden has an uncertainty avoidance index of 29 and the United Kingdom, 35. There is not much difference between their uncertainty avoidance indices. Yet, studies about the impact of anxiety on cesarean deliveries indicate two contradicting conclusions. Therefore, more studies are required about the impact of anxiety on cesarean deliveries.

CONCLUSIONS

We present a new perspective on factors that may impact cesarean delivery rates. Our study indicates that the masculinity index, representing gender differences in a country, has statistically significant influence on the cesarean rates. The risk avoidance index, a proxy for anxiety in a culture, affects cesarean rates negatively. However, studies performed on the impact of anxiety about the childbirth on cesarean deliveries reaches conflicting conclusions. Based on our studies, it is evident that culture does have an impact on the medical practice. Therefore, physicians and health care professionals should be made aware of the impact of culture on medical practice with a view to make medical practice more effective.

[For list of references, please contact the author]

A SIMULATION-BASED STUDY PLAN FOR PHARMACY MANAGEMENT

Esmail Mohebbi
Department of Management/MIS, University of West Florida, Pensacola, FL 32514
Ph: 850-863-6590, *emohebbi@uwf.edu*
Nooshin Abedini
P.O. Box 5274, *Navarre, FL* 32566
Ph: 850-313-6500, *n00shin@yahoo.com*

ABSTRACT

We consider the application of simulation modeling in pharmacy management. We identify a gap in the literature with regard to retail pharmacy management, and discuss the specific attributes which distinguish the operations of retail pharmacies from pharmacy operations in hospital settings. We then propose a plan for conducting a comprehensive simulation study of pharmacy management in a retail environment.

Keywords: Simulation modeling, pharmacy management, healthcare management, dynamic systems

INTRODUCTION

Simulation has long been recognized as an effective means of modeling and analysis of complex and dynamic systems where uncertainties and interdependencies of embedded events as well as conflicting objectives of stakeholders and decision makers pose difficult managerial challenges. It is therefore not surprising that a glance at the literature reveals a large volume of simulation applications in business and manufacturing sectors. To that end, simulation models continue to make contributions to disciplines such as science, business, engineering, entertainment, and defense [8]. More recently, the increasing concern over actual and/or potential problem areas in the healthcare sector around the world has created a new breeding ground for simulation models [6]. Healthcare delivery systems are routinely faced with the challenge of providing quality, affordable and easily accessible care to every member of the population that they are intended to serve. Growing populations, increased life expectancy, demographic changes, socio-economic developments, social welfare systems, shortage of healthcare professionals, cost containment, and government regulations are among the many factors that contribute to the complexity of such multi-disciplinary challenge [5].

Pharmaceutical expenditures are regarded as the biggest contributor to the rising health care costs across the nation [2]. While a large portion of such expenditures are due to the development of new drugs, increased usage, greater patent protection, and direct-to-consumer advertising; nevertheless, pharmacy management issues such as allocation of resources, inventory control, shortage of licensed pharmacists, scheduling of the pharmacy staff, streamlining of the workflow and the administrative processes, physical layout of the pharmacy, and automation of ordering and dispensing of medications play an important part in defining the overall productivity and cost efficiency of the pharmaceutical products delivery system. Hence, simulation modeling–owing to its proven success in process design/reengineering as well as improving system utilization in manufacturing, logistics, an customer-service applications–can be perceived as a viable tool for achieving improvement in pharmacy management. There are, however, only a few manifestations of simulation studies of pharmacy management in the literature.

This paper presents a review of existing simulation studies of pharmacy management and outlines a study plan for simulation modeling of a retail pharmacy operation management.

LITERATURE REVIEW

To the best of our knowledge, the earliest application of simulation modeling in pharmacy management appeared in Mukherjee [7]. The author presented a simulation study of operations in the pharmacy of the University of Tennessee hospital at Knoxville to analyze the activities in the night shift, and determine standard times for activities after computerization of work in the pharmacy–which had led to utilization of a computer database for maintaining patient profiles and generating labels for medications. The model accounted for the flow of regular, critical, and outdoor orders through the pharmacy; and focused on the turnaround times for theses orders and telephone calls as well as the utilization of the pharmacy staff (pharmacist, computer technician, and order technician) for the purpose of performance measurement. The experimental results included examining different scenarios for collecting regular orders, assigning pharmacists to various shifts, and prioritizing the different types of orders and phone calls. The study was concluded with a sensitivity analysis over input parameters such as the percentage of regular orders that need clarification from the prescribing physicians.

In Wong et al. [12], the authors applied simulation modeling to quantify and evaluate the advantages of utilizing an electronic medication ordering, dispensing and administration process compared with using a manual process in a hospital pharmacy setting at the Sunnybrook Campus of the Sunnybrook and Women's College Health Sciences Center located in Toronto, Ontario, Canada. Medication delivery failure rates, turnaround times, and resource requirements (physician, nurse, pharmacist, and order entering and dispensing pharmacy technicians) were chosen as performance measures for assessment purposes. Their comparative analysis showed considerable reductions in average turn-around times and failure rates in favor of the computerized system.

Spry and Lawley [10] developed a simulation model for inpatient pharmacy at the BroMenn Regional Medical Center in Normal, Illinois, to find the work schedule that helps the hospital keep up with a target turnaround time of 120 minutes for a prescription order. Their experimental results showed that adding evening staff (IV technicians) improves the turnaround time in the evening and into the early morning. The model also examined the timing of a daily medication cart exchange between the pharmacy and various areas in the hospital as well as the frequency of filling the machines that are located in the nurses' area and dispense commonly used and emergency medications. The results were not conclusive about the timing of the cart checking operation, but the authors suggested that the hospital should try to change to a once in the morning filling of the medication dispensing machines.

The objective of Yurtkuran and Emel [14] simulation study of a regional hospital pharmacy in Bursa, Turkey, was to minimize the turnaround time of the medication orders using the available resources. The study examined two alternative scenarios in comparison to the existing system. In particular, these scenarios were based on changing the starting time of preparing the daily medication packages at the pharmacy, and relaxing the timing restriction on order entries by the nurses and physicians. The experimental results showed that modifying the software of the hospital's computer system to allow for more flexibility in order entering, and transferring a technician and a pharmacist to the dayshift improve the system performance and decrease the turnaround time—resulting in a higher quality service.

Noting that hospital pharmacies in the US are experiencing inventory problems that result in waste, shortages, and delays for drug substitutions, Villa-Parrish et al. [11] studied inventory control in a hospital pharmacy. The objective of their study was to determine the inventory level for perishable pharmaceuticals that could minimize wastage and holding costs while maximizing timely access to the drugs. The development of the inventory control policy in their study was based on a Markov Decision Process model which accounted for demand as a function of patients' conditions. However, faced with computational intractability of the analytical model, the authors used simulation to evaluate two types of base-stocks inventory control policies.

Considering the above, it is clear that the current simulation studies of pharmacy management have solely focused on hospital pharmacy settings. The authors were unable to locate any documented study of retail pharmacy management in the literature. This paper proposes a plan for simulation modeling and analysis of retail pharmacy management.

RETAIL PHARMACY

Retail pharmacy is growing rapidly in the US. There are more than 39,000 pharmacies operated by traditional chain pharmacy companies, supermarkets and mass merchants. There are nearly 17,000 additional independent pharmacies operating across the country [9]. The Chain Drug Review [3] reports that the top drugstore chains by dollar volume include Walgreens ($53.76 billion), CVS ($45.09 billion) and Rite Aid ($24.33 billion). A related report [4] depicts the following distribution for the traditional chain drug stores merchandise mix in 2007: 71% prescription, 6.1% over the counter (OTC) medications, 5.2% groceries/consumables, 4.9% general/other merchandise, 4.2% toiletries, 3.3% non-edible consumables, 1.5% packaged alcohol and 1.2% vitamins/minerals.

At the same time, the expansion of retail pharmacies has contributed to the shortage of licensed pharmacists nationwide. According to the Bureau of Labor Statistics [1], '*pharmacists held about* 243,000 *jobs in* 2006*. About* 62 *percent worked in community pharmacies that were either independently owned or part of a drugstore chain, grocery store, department store, or mass merchandise….About* 23 *percent of pharmacists worked in hospitals. A small proportion worked in mail-order and Internet pharmacies, pharmaceutical wholesalers, offices of physicians, and the Federal Government*.' The same source predicts that demand for licensed pharmacists is expected to grow by 22% between 2006 and 2016.

Considering the above, devising effective means of retail pharmacy management in areas such as work-force scheduling, prescription management and patient consultation appear to be critical in providing the consumer with quality care. To that end, such managerial practices should explicitly account for operational characteristics that distinguish the job of a pharmacist in a retail store from that of a pharmacist in hospital settings. Pharmacists in community pharmacies are known for performing the following tasks [1]:

- Dispense medications
- Counsel patients on the use of prescription
- Counsel patients on the use over-the-counter medications
- Advise physicians about patients' medication therapy
- Advise patients about general health topics such as diet, exercise, and stress management
- Provide patients with information on products such as durable medical equipment or home health care supplies
- Deal with insurance processing issues
- Sell non-health-related merchandise
- Provide specialized services to help patients with conditions such as diabetes, asthma, smoking cessation, or high blood pressure, administer vaccinations
- Hire and supervise personnel, and oversee the general operation of the pharmacy

Other activities include seeking clarifications on medications from prescribing physicians, transferring prescription (electronically or through phone/fax) to/from other pharmacies at the request of patients, inventory control and ordering of drugs and controlled substances, performing drug utility review, compounding medications, selling ephedrine-contained OTC medications, completing on-line review courses, dealing with drive-through customers, and resolving customer service problems related to promotional offers, discounts and coupons. Clearly, while the degree of direct involvement of a pharmacist in all these categories may vary from one drugstore to another, the fact remains that interacting with patients places a significant demand on the pharmacist's time and thereby directly impacts the overall performance of the pharmacy operation. As noted by the Bureau of Labor Statistics [1], '*pharmacists are becoming more involved in counseling patients and planning drug therapy programs*.' By contrast, pharmacists in health care facilities have limited direct contact (if any) with the patients. For example, they may provide counseling to hospitalized patients on the use of medication at home before they get discharged [1]. Hence, the applicability of the existing simulation models of hospital pharmacy operations to retail settings is questionable at best since they do not adequately reflect the interaction with patients as the main source of uncertainty in the modeling process.

This study plan is aimed at addressing the existing gap in the literature by developing a comprehensive simulation model of pharmacy operations in a retail environment. The next section presents an outline of our plan.

STUDY PLAN

Objective

Our objective is to develop and utilize a simulation-based decision support tool to examine various alternatives for work-force scheduling, and process flow of operations in a busy 24-hour drugstore.

Model development and data collection

The development, verification and validation of the simulation model will be based on the process-flow mapping and data of a major drugstore chain in southeast of the US. Our preliminary contacts have been directed toward a non-stop 24-hour pharmacy with drive-through facility to broaden the applicability of the simulation results.

Performance measures

Our conversations with numerous retail pharmacists have led us to selecting the following measures for performance assessment purposes:

- Turn-around time for called-in (including fax and email) new prescriptions
- Turn-around time for called-in refill orders
- Turn-around time for waiting (i.e., drop-offs) new prescriptions
- Turn-around time for waiting refill prescriptions
- Turn-around time for in-person patient consultations
- Turn-around time for phone calls
- Waiting time for drive-through customer
- Waiting time for OTC-related inquiries
- Delay (defined as deviation from the promised or requested time of order completion) in filling called-in new prescriptions
- Delay in processing called-in refill orders
- Delay in filling waiting new prescriptions
- Delay in processing waiting refill orders
- Utilization of pharmacist

Random Variables
The model will take into account the following variables:

- Inter-arrival times of called-in prescriptions
- Inter-arrival times of refill orders
- Inter-arrival times of waiting prescriptions
- Inter-arrival times of waiting refills
- Inter-arrival times of drive-through customers
- Inter-arrival times of phone calls
- Inter-arrival time of OTC-related inquiries
- Inter-arrival times of general merchandise inquiries
- New order entry time
- Refill order entry time
- Likelihood of seeking clarification from prescribing physicians
- Likelihood of seeking clarification from the patient
- Preliminary order checking time for new prescription
- Preliminary order checking time for refills
- New prescription order filling time
- Refill order filling time
- Secondary checking time for new prescriptions
- Secondary checking time for refills
- Likelihood of seeking clarification from the insurance company
- Response time to OTC-related inquiries
- Response time to general merchandise inquiries

Software
The model will be developed within the framework of the Simul8 simulation software. Data analysis will be conducted using Stat:Fit.

Simulation runs
The data will be collected for one full week of operations at the designated pharmacy. The model will be run for multiple replications of year-long operations (52 weeks).

REFERENCES

[1] http://www.bls.gov/oco/ocos079.htm (Bureau of Labor Statistics)
[2] http://www.bluecrossmn.com/bc/wcs/groups/bcbsmn/@mbc_bluecrossmn/documents/public/mbc1_hl_drugs_story.hcsp (Blue Cross and Blue Shield of Minnesota)
[3] http://www.chaindrugreviewmediakit.com/r5/showkiosk.asp?listing_id=1749201 (Chain Drug Review)
[4] http://www.chaindrugreviewmediakit.com/r5/showkiosk.asp?listing_id=1749586 (Chain Drug Review)
[5] Eldabi, T., Young, T., Toward a framework for healthcare simulation, Proceedings of the 2007 Winter Simulation Conference, 2007, 1454-1460.
[6] Kujis, J., Paul, R.J., Stergioulas, L.K., Can healthcare benefit from simulation methods in the same way as business and manufacturing has? Proceedings of the 2007 Winter Simulation Conference, 2007, 1449-1453.
[7] Mukherjee, A.K., A Simulation model for management of operations in the pharmacy of a hospital, Simulation, 56 (1991) 91-103.

[8] National Research Council, Modeling and simulation in Manufacturing and Defense Systems, National Academic Press, Washington DC, 2002.
[9] http://quezi.com/4157
[10] Spry C.W., Lawley, M.A., Evaluating hospital pharmacy setting and work scheduling using simulation, Proceedings of the 2005 Winter Simulation Conference, 2005, 2256-2263.
[11] Villa-Parrish, A.R., Ivy, J.S., King, R.E., A simulation-based approach for inventory modeling of perishable pharmaceuticals, Proceedings of the 2008 Winter Simulation Conference, 2008, 1532-1538.
[12] Wong, C., Geiger, G., Derman, Y.D., Rusby, C.R., Carter, M.W., Redesigning the medication ordering, dispensing, and administration process in an acute academic health sciences center, Proceedings of the 2003 Winter Simulation Conference, 2003, 1894-1902.
[13] Workman, R.W., Simulation of the drug development process: A case study from the pharmaceutical industry, Proceedings of the 2000 Winter Simulation Conference, 2000, 1995-1998.
[14] Yurtkuran, A., Emel, E., Simulation based decision-making for hospital pharmacy management, Proceedings of the 2008 Winter Simulation Conference, 2008, 1539-1546.

THE EXCHANGE RATIO IN THE DIMENSIONS OF INTEGRITY, CONFIDENTIALITY AND AVAILABILITY OF INFORMATION SECURITY IN THE TEACHING AND GENERAL HOSPITALS: CASE STUDY OF MINISTRY OF HEALTH HOSPITALS OF TURKEY

Menderes TARCAN: Ph.D., Ministry of Health, Performance Management and Quality Improvement Directorate, Ankara, Turkey. E-mail:mtarcan@hacettepe.edu.tr, Phone: +90 533 307 51 32.
Yener GUL: MsC., Ministry of Health, Performance Management and Quality Improvement Directorate, Ankara, Turkey. E-mail: kalitegelistirme@hotmail.com, Phone: +90 535 899 56 96.
Fatma GUL: MsC., Middle East Technical University, Institute of Applied Sciences, Ankara, Turkey. E-mail:gulfatmagul@gmail.com, Phone: +90 506 323 90 48.
Gamze Y. TARCAN: MsC., Hacettepe University, Health Services Technical School, Ankara, Turkey. E-mail:gamze@hacettepe.edu.tr, Phone: +90 538 825 88 15.

ABSTRACT

The aim of this study is to determine the level of information security in hospitals which are education-research and public service hospitals of Ministry of Health of the Republic of Turkey by measuring information securities during 2008 and 2009.

In this context, we did check up on the three basic factors forming "the dimensions of information security". These three factors are integrity, confidentiality and availability. The overall success ratio of the Information Security during 2008 and 2009 as follows: in 2008, while the overall success of information security is 93.28 % in the Ministry of Health hospitals (education-research hospitals and public service hospitals), this ratio has increased to 93.82 % in 2009.

Key words: information security, dimensions of the information security in hospitals, hospital information infrastructure.

Introduction

Nowadays, security attacks against informatic systems and processed data by these systems increases rapidly. These attacks intended for computer systems and networks can cause loss of money, time, prestige and valuable information. If these attacks are directed to systems which affects life, in this situation, the loss can be human life [4] [5].

Information security in hospitals, is a process that begins with the description of risks which threats the confidentiality, integrity and availability of inventory information in hospitals. Therefore risk management requirements must be fulfilled. At this point, information security is a total of processes and technologies that are used to prevent disclosure, abuse and unauthorized access of valuable information [6] [7]. According to this, information security can be analyzed in three basic dimensions as in Figure 1.

Figure 1: Confidentiality, integrity and availability triad

Integrity can be defined as the determination of accuracy and integrity of information and methods related with information processing [2]. In this study, following criteria related with dimensions of integrity in information security is asked; *"A database that can be accessed only by the system administrator and with read-only property must exists in the hospital automation system which holds the records of users that log in to the system, transactions done by users, changes on the system settings and system errors"*.

Confidentiality is the guarantee that information is accessible only to persons with access permission. Confidentiality means that special information in system does not get out of [1]. The following criteria is in our confidentiality content; *"Hospital should have a software that enables persons who have formerly permission to access records and operations of services regarding patients"*.

Availability is the situation that systems serve non-stop and information in these systems do not get lost and they serve always availability of acces. In our survey, we have asked the following criteria which is related with this topic; *"the backup of data in hospital information system should be done orderly in another storage excluding server"*.

An information system can be accepted as safe if it provides confidentiality, integrity and availability so-called CIA triad [3]. The questioning of three basic dimensions of information security and the comparison of their states in hospitals form the main purpose of this study.

Material and Method

The content of this study includes 2nd and 3rd level hospitals of Ministry of Health. No samples are selected for this study. Total field under survey is whole 2nd and 3rd level hospitals of Ministry of Health excluding district hospitals and health centers. "Service Quality Standards Scale" developed by the Ministry of Health was used to evaluate the institutional infrastructure of hospitals. Data obtained from implementation of this scale was assessed by frequency analysis.

Service Quality Standards Scale, developed by the Ministry of Health in 2007, was used to measure the success level of hospitals about the information security. This scale is a scale that hospitals are evaulated under 21 areas such as Clinics, Laboratories, Imaging, Surgical Units, Clinics, Intensive Care Units, Dialysis Services, Emergency Services, Pharmacy Services, Patient and Employee Safety, Corporate Services Management, Hospital Information System, Facility Management and Security, etc. In this measurement tool, hospital scores “5” points if it is able to meet the requirements, if not scores “0” points. Therefore as the scores approach to “5”, the success of hospital in the information security is apparent, on the other hand if scores approach to “0”, the success of hospital in the information security is in the lowest level. The capacity of each hospital's success in information security was calculated by the evalution of integrity, confidentiality and availability areas.

Table 1 shows the capacity of integrity, confidentiality and availability of education-research hospitals that operate under Ministry Of Health for 2008. In 2008, overall succes rate of information security has been calculated for 56 hospitals. They have 439 beds in average. Success percentage of information security in the Ministry of Health hospitals (education-research hospitals) has been found as 96.43 %. When hospitals are qualified according to dimensions form information security, scores for integrity, confidentiality and availability was determined as 5.00, 4.46 and 5.00, respectively. It is observed that the capacity of confidentiality dimension of information security was lower than the capacity of integrity and availability dimensions of information security in education-research hospitals. Also when an assessment is made according to the types of education-research hospitals, Cardiovascular Surgery Hospitals are within the group that has the lowest score .

Table 1: The capacity of integrity, confidentiality and availability of training hospitals of Ministry of Health in the scope of information security (2008)

Type of Hospital	N	Beds	Integrity	Confidentiality	Availability	Success Rate %
T., General	31	508	5.00	4.35	5.00	95.70
T., Cardiac&Cardiovascular Surgery	4	284	5.00	3.75	5.00	91.67
T., Thorax&Thorax surgery	4	521	5.00	5.00	5.00	100.00
T., Gynecology-Obstetrics-Pediatrics	1	540	5.00	5.00	5.00	100.00
T., Gynecology-Obstetrics	6	339	5.00	4.17	5.00	94.44
T., Pediatrics	2	316	5.00	5.00	5.00	100.00
T., Physiotherapy	2	267	5.00	5.00	5.00	100.00
T., Skeletal&Bone Diseases	1	133	5.00	5.00	5.00	100.00
T., Psychiatry	2	549	5.00	5.00	5.00	100.00
T., Ophthalmology	2	94	5.00	5.00	5.00	100.00
T., Oncology	1	416	5.00	5.00	5.00	100.00
Average	**56**	**439**	**5.00**	**4.46**	**5.00**	**96.43**

Table 2 shows the capacity of integrity, confidentiality and availability of general service hospitals that operate under Ministry Of Health. In 2008, overall succes rate of information security has been calculated for 495 hospitals. They have 124 beds in average. Success percentage of information security in these hospitals has been found as 92.93 %. When hospitals are qualified in terms of dimensions form information security, scores for integrity, confidentiality and availability was determined as 4.77, 4.49 and 4.68, respectively. It is observed that the capacity of confidentiality dimension of information security was lower than the capacity of integrity and availability dimensions of information security in education-research hospitals. Also when an assessment is made according to the types of general service hospitals, Dermal, Leprosy and Venereal Diseases Hospitals are within the group that has the lowest score.

Table 2: The capacity of integrity, confidentiality and availability of General Service Hospitals of Ministry of Health in the scope of information security (2008)

Type of Hospital	N	Beds	Integrity	Confidentiality	Availability	Success Rate %
Beds 0-49	128	29	4.80	4.65	4.53	93.23
Beds 50-99	130	56	4.81	4.58	4.81	94.62
Beds 100-499	149	196	4.77	4.33	4.77	92.39
Beds 500 and more	19	506	5.00	4.47	4.74	94.74
Gynecology-Obstetrics-Pediatrics	29	159	4.48	4.31	4.66	89.66
Gynecology-Obstetrics	2	209	5.00	5.00	5.00	100.00
Pediatrics	4	215	3.75	3.75	5.00	83.33
Physiotherapy	7	134	5.00	5.00	5.00	100.00
Skeletal&Bone Diseases	1	240	5.00	5.00	5.00	100.00
Cardiac&Cardiovascular Surgery	1	400	5.00	5.00	5.00	100.00
Thorax&Thorax surgery	11	131	4.55	5.00	4.55	93.94
Psychiatry	5	437	5.00	5.00	4.00	93.33
Dermal, leprosy&venereal diseases	3	65	3.33	1.67	1.67	44.44
Emergency&Traumatology	4	63	5.00	5.00	5.00	100.00
New	2	35	5.00	2.50	2.50	66.67
Average	**495**	**124**	**4.77**	**4.49**	**4.68**	**92.93**

Table 3 shows the capacity of integrity, confidentiality and availability of all hospitals that operate under Ministry Of Health for 2008. In 2008, overall success rate of information security has been calculated for 551 hospitals. They have 156 beds in average. Overall success of information security in the Ministry of Health hospitals (education-research hospitals and public service hospitals) has been found as 93.28%. However when hospitals are qualified according to dimensions form information security, scores for integrity, confidentiality and availability was determined as 4.79, 4.49 and 4.71, respectively.

Table 3: The capacity of integrity, confidentiality and availability of Ministry of Health hospitals in the scope of information security (2008)

Type of Hospital	N	Beds	Integrity	Confidentiality	Availability	Success Rate %
T., General	31	508	5.00	4.35	5.00	95.70
T., Cardiac&Cardiovascular Surgery	4	284	5.00	3.75	5.00	91.67
T., Thorax&Thorax surgery	4	521	5.00	5.00	5.00	100.00
T., Gynecology-Obstetrics-Pediatrics	1	540	5.00	5.00	5.00	100.00
T., Gynecology-Obstetrics	6	339	5.00	4.17	5.00	94.44
T., Pediatrics	2	316	5.00	5.00	5.00	100.00
T., Physiotherapy	2	267	5.00	5.00	5.00	100.00
T., Skeletal&Bone Diseases	1	133	5.00	5.00	5.00	100.00
T., Psychiatry	2	549	5.00	5.00	5.00	100.00
T., Ophthalmology	2	94	5.00	5.00	5.00	100.00
T., Oncology	1	416	5.00	5.00	5.00	100.00
Beds 0-49	128	29	4.80	4.65	4.53	93.23
Beds 50-99	130	56	4.81	4.58	4.81	94.62
Beds 100-499	149	196	4.77	4.33	4.77	92.39
Beds 500 and more	19	506	5.00	4.47	4.74	94.74
Gynecology-Obstetrics-Pediatrics	29	159	4.48	4.31	4.66	89.66
Gynecology-Obstetrics	2	209	5.00	5.00	5.00	100.00
Pediatrics	4	215	3.75	3.75	5.00	83.33
Physiotherapy	7	134	5.00	5.00	5.00	100.00
Skeletal&Bone Diseases	1	240	5.00	5.00	5.00	100.00
Cardiac&Cardiovascular Surgery	1	400	5.00	5.00	5.00	100.00
Thorax&Thorax surgery	11	131	4.55	5.00	4.55	93.94
Psychiatry	5	437	5.00	5.00	4.00	93.33
Dermal, leprosy&venereal diseases	3	65	3.33	1.67	1.67	44.44
Emergency&Traumatology	4	63	5.00	5.00	5.00	100.00
New	2	35	5.00	2.50	2.50	66.67
Average	**551**	**156**	**4.79**	**4.49**	**4.71**	**93.28**

Table 4 shows the capacity of integrity, confidentiality and availability of education-research hospitals that operate under Ministry Of Health for 2009 year. In 2008, overall succes rate of information security has been calculated for 54 hospitals. They have 494 beds in average. Success percentage of information security in the Ministry of Health hospitals (education-research hospitals) has been found as 90.74 %. When hospitals are qualified according to dimensions form information security, scores for integrity, confidentiality and availability was determined as 4.84, 4.07 and 4.72, respectively. It is observed that the capacity of confidentiality dimension of information security was lower than the capacity of integrity and availability dimensions of information security in education-research hospitals. Also when an assessment is made according to the types of education-research hospitals, Cardiovascular Surgery hospitals are within the group that has the lowest score, same as in 2008. To find the reasons of this result, it was made an interview with hospital managers. Consequently, the following conclusions were found;

- During the research, information systems and software development effort were in progress and,

- During this survey period, the software belong to hospital information system is in the process of replacement in some of these hospitals (2 of 4 hospitals)

Table 4: The capacity of integrity, confidentiality and availability of training hospitals of Ministry of Health in the scope of information security (2009)

Type of Hospital	N	Beds	Integrity	Confidentiality	Availability	Success Rate %
T., General	28	604	5.00	4.29	4.82	94.05
T., Cardiac&Cardiovascular Surgery	5	363	3.00	3.00	3.00	60.00
T., Thorax&Thorax surgery	4	549	5.00	3.75	5.00	91.67
T., Gynecology-Obstetrics-Pediatrics	1	460	5.00	5.00	5.00	100.00
T., Gynecology-Obstetrics	6	333	5.00	5.00	5.00	100.00
T., Pediatrics	2	401	5.00	5.00	5.00	100.00
T., Physiotherapy	2	205	5.00	5.00	5.00	100.00
T., Skeletal&Bone Diseases	1	133	5.00	5.00	5.00	100.00
T., Psychiatry	2	579	5.00	2.50	5.00	83.33
T., Ophthalmology	2	101	5.00	2.50	5.00	83.33
T., Oncology	1	607	5.00	0.00	5.00	66.67
Average	**54**	**494**	**4.81**	**4.07**	**4.72**	**90.74**

Table 5 shows the capacity of privacy and availability of the hospitals within the scope of information security, that operate under Ministry Of Health in 2009. In 2009, the success percentage points in information security of 469 hospitals were calculated. These hospitals have 138 beds in average. The average information securty point of the hospitals is 94.17%. If the hospitals were evaluated in seperate terms that forms information security; integrity score (4.84), confidentiality score (4.50) and the usability score (4.79) were identified. We can see that the score of confidentiality is under the scores of integrity and usability. In addition, if the general assessment is made according to the type of hospital services, the bone disease hospitals constitute the least succesfull group in the success of information security.

Table 5: The capacity of integrity, confidentiality and availability of General Service Hospitals of Ministry of Health in the scope of information security (2009)

Type of Hospital	N	Beds	Integrity	Confidentiality	Availability	Success Rate %
Beds 0-49	121	28	4.83	4.50	4.75	93.94
Beds 50-99	124	62	4.84	4.44	4.76	93.55
Beds 100-499	145	215	4.79	4.48	4.86	94.25
Beds 500 and more	20	567	5.00	4.50	4.75	95.00
Gynecology-Obstetrics-Pediatrics	27	170	5.00	4.63	4.81	96.30
Gynecology-Obstetrics	2	162	5.00	5.00	5.00	100.00
Pediatrics	4	240	5.00	5.00	5.00	100.00
Physiotherapy	6	138	5.00	5.00	5.00	100.00
Skeletal&Bone Diseases	2	251	5.00	2.50	5.00	83.33
Cardiac&Cardiovascular Surgery	1	411	5.00	5.00	5.00	100.00
Thorax&Thorax surgery	7	138	5.00	4.29	4.29	90.48
Psychiatry	4	512	5.00	5.00	3.75	91.67
Dermal, leprosy&venereal diseases	3	85	5.00	5.00	5.00	100.00
Emergency&Traumatology	3	62	3.33	5.00	5.00	88.89
Average	**469**	**138**	**4.84**	**4.50**	**4.79**	**94.17**

Table 6 shows the capacity of integrity, confidentiality and availability of all hospitals that operate under Ministry Of Health. In 2009, overall succes rate of information security has been calculated for 523 hospitals. They have 175 beds in average. Overall success of information security in the Ministry of Health hospitals (education-research hospitals and public service hospitals) has been found as 93.82%. However when hospitals are qualified according to dimensions form information security, scores for integrity, confidentiality and availability was determined as 4.84, 4.46 and 4.78, respectively.

Table 6: The capacity of integrity, confidentiality and availability of Ministry of Health hospitals in the scope of information security (2009)

Type of Hospital	N	Beds	Integrity	Confidentiality	Availability	Success Rate %
T., General	28	604	5.00	4.29	4.82	94.05
T., Cardiac&Cardiovascular Surgery	5	363	3.00	3.00	3.00	60.00
T., Thorax&Thorax surgery	4	549	5.00	3.75	5.00	91.67
T., Gynecology-Obstetrics-Pediatrics	1	460	5.00	5.00	5.00	100.00
T., Gynecology-Obstetrics	6	333	5.00	5.00	5.00	100.00
T., Pediatrics	2	401	5.00	5.00	5.00	100.00
T., Physiotherapy	2	205	5.00	5.00	5.00	100.00
T., Skeletal&Bone Diseases	1	133	5.00	5.00	5.00	100.00
T., Psychiatry	2	579	5.00	2.50	5.00	83.33
T., Ophthalmology	2	101	5.00	2.50	5.00	83.33
T., Oncology	1	607	5.00	0.00	5.00	66.67
Beds 0-49	121	28	4.83	4.50	4.75	93.94
Beds 50-99	124	62	4.84	4.44	4.76	93.55
Beds 100-499	145	215	4.79	4.48	4.86	94.25
Beds 500 and more	20	567	5.00	4.50	4.75	95.00
Gynecology-Obstetrics-Pediatrics	27	170	5.00	4.63	4.81	96.30
Gynecology-Obstetrics	2	162	5.00	5.00	5.00	100.00
Pediatrics	4	240	5.00	5.00	5.00	100.00
Physiotherapy	6	138	5.00	5.00	5.00	100.00
Skeletal&Bone Diseases	2	251	5.00	2.50	5.00	83.33
Cardiac&Cardiovascular Surgery	1	411	5.00	5.00	5.00	100.00
Thorax&Thorax surgery	7	138	5.00	4.29	4.29	90.48
Psychiatry	4	512	5.00	5.00	3.75	91.67
Dermal, leprosy&venereal diseases	3	85	5.00	5.00	5.00	100.00
Emergency&Traumatology	3	62	3.33	5.00	5.00	88.89
Average	**523**	**175**	**4.84**	**4.46**	**4.78**	**93.82**

Conclusion and Recommendations

Due to the variety of roles and types of service, accessibility of information needed in the planning of service production and production mamagement requires a rather complex and difficult process. Therefore, information production and methods for accessibility of information should be designed properly from the first step. In the study, "Hospital Success Rates of Information Security" were higher both in 2008 and 2009. However, the level of knowledge of managers and employees who work on "Information Security and Hospital Information Systems Management" must have studied. To keep the data safe from the attacks within the hospital network or outside the network, there has to be a highly reliable automation system which runs non-stop on a 24 hours basis.

References

[1] Hunter, M.D. *An Information Security Handbook,* Springer-Verlag, London, 2001.

[2] Alberts, C. and Dorofee, A. *Managing Information Security Risks: The Octave Approach.* Pearson Education Inc., 2004.
[3] Blanas, G. Strategic Open Software Development Illness, 2003, <http://opensource.mit.edu/papers/blanas1.pdf>. Accessed 29 October, 2008.
[4] Ministry of Health, *Information Security Act*, 2005.
[5] The Prime Minister's Office, *Patient Rights Act*, 1998.
[6] Furnell S. *IFIP Workshop-Information Security Culture,* Workshop, Plymouth University, 2006.
[7] Ogut P. *Policy on Information Security in the Global World*, Ankara University, PhD. Dissertation, 2006.

JOB MARKET PERCEIVED BY COLLEGE GRADUATES

Margaretha Hsu, Shippensburg University, Shippensburg, Pennsylvania 17257, 717-477-1422
mmhsu@ship.edu

ABSTRACT

This paper will focus on the job searching experience college graduates in business have. It will also explore the impact that field of study, grade point average, internship, personal connections, and gender have on the perceptions of job market these graduates indicated. The relationship between when the job search begins and the length of time for graduate to locate an employment is also examined.

Keywords: job market, field of study, internship

INTRODUCTION

It is widely believed that the more education you have the more you will earn and the less risk you have of being unemployed. However, Mandel [3] reported on Business Week that college graduates' earnings have fallen since year 2000. Technology, globalization, and economy may also cause the job markets for college graduates to shrink. The problems that college graduates face in job searching are complex. They may vary widely based on major field of study, individual aptitude, personal circumstances, and geographic location. Schomburg [4] confirms the common view that the fields of study play a major role in graduates' occupational prospects. College graduates from certain fields of study are more likely to face employment challenges. Freeman, Snyder & Connolly [2] have found that there is a significant variation in salaries due to differences in gender and college major selection. Study by Torpe [5] suggests that the field of study also makes a difference in how quickly college graduates settle into jobs that require their degree. Dohm & Wyatt [1] find that experience contributes to salary variations within occupations. For many college students, internship is considered their most valuable work experience, with an expectation that internship will give them a competitive advantage in the workplace. The objective of this study is to investigate whether the perceptions of job market by business graduates are related to elements such as, their areas of study, when job search begins, how quickly a job if sound, GPA, personal connections, gender, and internship experience.

METHODOLOGY

A basic information-gathering instrument was developed and distributed to all college of business graduates from a local university two years after graduation. Data on the length of time to secure an employment, various job search experiences, the time in which the job search began, work experience prior to graduation, major, GPA, and gender were collected for the class of 2002 to the class of 2006 two years after graduation. A total of 754 data observations are obtained for the study.

RESULTS

During the 2002 to 2006 time period, approximately half the graduates in our survey reported that job searching was more difficult than expected. Moreover, a significantly large percentage of students who started searching for a position after graduation indicated this sentiment more than those who started job searching before graduation. As expected, the field of study affected graduate's job searching experience. It was found that over the five year period, between 2002 and 2006, in terms of the percentage of graduates that felt job hunting was more difficult than expected, accounting majors had the lowest with an average of 30%, followed by finance majors with an average of 40%; they were also mostly stable across 5-year period. Marketing majors had an average of 58% and management majors had an average of 62%, and they were also quite stable over the same 5 year period. The percentages for decision science majors and business information system majors fluctuated from year to year.

The source of position opening information used by our students was also investigated. The survey data showed that 24.1% found their position through internet, followed by connections through families and friends (23.5%), internship (11.9%), University Career Development Center (9.4%), newspaper (8.2%), job fairs (4.6%), and others.

Chi-square test confirms the notion that the sooner the college students start their job search the better chance they have in locating one. Data showed that among those who started job searching before graduation, 50% of them secured a position before graduation and only 7% took longer than six months to locate an employment in their field of study. On the other hand, for students who started looking for a job after graduation, 28% of them took longer than six months to find a job.

There was a slight increase in the percentage of students complete the internship from 2002 to 2006. Among students who had completed internship, 73% of them indicated that internship was a very importance experience.

In an attempt to study the perceived importance of GPA, internships, extra-curricular activities, leadership activities, core courses, and course work in one's major in job searching, the data was divided into two groups. Group 1 consisted of those who have been employed full time in their field of study or in related fields, and group 2 consisted of the remaining respondents. Several Chi-square tests were conducted. The results were somewhat surprising. The only significant difference in opinions between the two groups was in the area of course work in their major field. Respondents in group 2 did not view their major course work was as important as group 1 in their job searching.

CONCLUSION

There are many factors related to college students' job prospect. Many students have experienced more problems in job searching than what they have anticipated. This study shows that college students will benefit from starting job searching earlier. It not only reduces the stress, but also helps to locate a job faster. Having an internship experience will definitely strengthen college student's resume. Internet is the most popular source of information on job openings. However, personal connections are also an important source of information. In addition, field of study do have an impact on student's job searching experience, but, gender is not a significant factor for business graduates.

REFERENCES

[1] Dohm, Arlene, Wyatt, Ian. " College at work: Outlook and earnings for college graduates, 2000-2010." *Occupational Outlook Quarterly*, Fall 2002.

[2] Freeman, Catherine, Snyder, Thomas. Connolly, Brooke. "The impact of degree field on the earnings of male and female college graduates." Annual meeting of American Educational Research Association, 2004. April, San Diego, California.

[3] Mandel, Michael. "College: the payoff shrinks." Business Week, September 12, 2005.

[4] Schomburg, Harald. "The professional success of higher education graduates." *European Journal of Education*, 2007, Vol. 42, No.1.

[5] Torpe, Elka Maria. "Earnings and occupations by college major, 1 and 10 years after graduation." *Occupational Outlook Quarterly*, Summer 2008.

PERSPECTIVES ON RESEARCH METHODS USED IN E-GOVERNMENT

Rhoda C. Joseph, Pennsylvania State University Harrisburg, (717) 948-6144, ruj1@psu.edu

ABSTRACT

The field of e-egovernment represents a unique microcosm in the universe of information systems. The field of e-government is strategically positioned across multi-disciplines including information systems, computer science, management and public administration. The interdisciplinary nature of the field of e-government warrants a purposive look at the variety of methodological opportunities available. Broadly defined e-government involves the use of information and communications technology to facilitate government interaction with citizens, employees, businesses and other governments. This *work-in-progress* paper presents a theoretical framework for examining the methodological opportunities available in the field of e-government.

KEYWORDS

e-government, electronic government, G2C, G2G, G2B, G2E, methods

INTRODUCTION

E-government facilitates communication between a government and its citizens via web-enabled computer technologies [5]. E-government also involves the use of technology to facilitate and support governmental interactions with non-citizen stakeholders including employees, businesses, and other government agencies. E-government projects are differentiated into categories based on the stakeholder involved. The four main categories of e-government are government to citizen (G2C), government to business (G2B), government to employee (G2E), and government to government (G2G). Each of the above categories highlights the specific stakeholder that interacts with a government.

E-government is an ideal field for the application of mixed methodology projects, which combine quantitative and qualitative research approaches in the same study. Mixed method studies can generate more rigorous insights and richer explanations and interpretations than the findings of a study relying on only a quantitative or qualitative method [2]. Quantitative studies are primarily geared at hypothesis testing and generally relying on surveys and experiments; qualitative studies use methods designed to develop theories and involve techniques such as interviews and case studies. The extant literature indicates that both quantitative and qualitative perspectives have been adopted in e-government research [9]. However, researchers tend to be proficient in, and philosophically lean more towards one methodology over another.

Should quantitative and qualitative methods be considered polar opposites? Hardly so, in the middle ground are advocates who argue that mixed methods can potentially harness the benefits provided by each perspective. While quantitative research approaches seek to generalize results, qualitative approaches seek to provide a richer understanding of a concept or event. Arguably, e-government research provides a unique opportunity to create synergy between quantitative and qualitative methods because of the field's interdisciplinary nature; practice areas; practical domains; and social, economic, and political implications. The collective diversity of e-government spans multiple dimensions including conceptual frameworks, evaluation techniques and theoretical approaches [10].

The mid 1990's witnessed the dawn of e-government, essentially driven by the growth and development of web-based technologies and electronic commerce. Countries such as the United States, the United Kingdom, and Australia have lead the way in facilitating more immediate and effective communication between the government and its constituents through e-government applications [9]. As a rich emerging area for research and practice, e-government can benefit from the application of theories, concepts, and methods applied and developed elsewhere in the information systems field. Further, investigating e-government via mixed-method studies can enhance the richness of questions and interpretations sought and encountered in e-government research.

THEORETICAL BACKGROUND

The principles of quantitative research are based on the hypothetico-deductive method and consist of seven main steps [13]:

1. *Observation:* Examine a phenomenon or situation and identify a research question(s), for example, who are the main users of e-government websites?
2. *Preliminary information gathering:* Conduct a literature review of existing papers, articles, and reports to try to find current and prevailing answers to your question(s).
3. *Theory formulation:* Identify or develop a theory to explain the phenomenon or situation you observed.
4. *Hypothesizing:* Write a set of specific propositions or hypotheses about the phenomenon that will be tested.
5. *Data collection:* Distribute a survey or run an experiment to test hypotheses.
6. *Data analysis:* Use appropriate statistical techniques to determine support for any of the stated hypotheses.
7. *Deduction:* Make conclusions and inferences based on the outcomes of data analysis.

Researchers that subscribe to the quantitative philosophy use empirical data collection techniques such as surveys, experiments, simulations, and modeling to test hypotheses, analyze results, and explain and generalize findings. Survey research has provided important insights into citizen usage of e-government technology. For example, a national public opinion survey conducted in the United States indicated that e-government has the potential to drive positive changes between a government and its citizens by increasing information delivery and responsiveness [16].

Further, quantitative researchers use random sampling with a large sample size to reduce sampling errors, while qualitative researchers use purposeful sampling with a relatively small number of participants [3]. Quantitative researchers generally strive to be impartial and report on fact-based hypothesis testing. Generalizability (ability to apply results to other samples and settings) and replicability (ability to repeat tests) are two hallmarks of the quantitative method, which generally seeks to explain and empirically establish cause-and-effect relationships. The desire for factual representation and a reliance on statistical significance to explain events form the basis for some of the criticisms that have surfaced towards the quantitative method. Its critics argue that the world does not exist merely in the mathematical realm and that not all explanations can be derived from hypothesis testing.

Qualitative research, on the other hand, involves data collection via methods such as interviews, participant observations, ethnography (examination of social behaviors, practices and norms), and case studies with explicit goals of using rich data to understand and explain phenomena. A recent case study explored e-government project management in Singapore illustrating a chronology of developments that occurred [8]. One additional study in the e-government literature utilized a qualitative meta-synthesis methodology to develop a comprehensive model of e-government development stages [15]. The meta-synthesis method involved the analysis of multiple e-government developmental studies, resulting in a single model that captures the various stages of e-government growth. The above two studies represent a short sample of the value of qualitative research in the field of e-government.

Qualitative research methods are widely accepted and developed in a variety of fields including management, sociology, psychology, and anthropology. Critics of the qualitative method argue that it can be subjective and generally fails to produce generalizable findings. However, subjectivity is part of the richness derived through qualitative methods and qualitative research, by definition, does not intend to generalize its findings but rather to explain specific phenomena in much greater depth. The qualitative methodology uses data that is rich and naturally occurring and involves iterative processes of data collection, data reduction, conclusion formation, and result verification [11].

Researchers will inevitably have a methodology that they prefer to adhere to stemming from skill proficiency, philosophical orientation, socialization through doctoral research, and other environmental influences. Nevertheless, researchers can be experts in one methodology but also appreciate the values of other orientations. Combining quantitative and qualitative approaches provides an opportunity for intriguing scholarly work in the domain of e-government.

RESEARCH METHODS IN E-GOVERNMENT

Extant literature in the field of e-government covers a wide selection of topics, levels, towns, cities, countries, regions and continents across the globe. A study of 110 peer reviewed articles in the e-government literature found that 24% of the studies focused on conceptualization, 18% on technology diffusion, 34% on e-service, and 25% on e-democracy [1]. A more recent study further showed that a majority of e-government studies are either conceptual in nature, or are case studies [4]. This indicates that almost one quarter of e-government studies have been theoretical papers. The past focus on theory building papers might be directly related to the youthfulness of the e-government field. A cursory look at e-government published articles also finds a fair subset of case studies and survey papers. The period 1998-2008 represents the foundation years of the field of e-government. In terms of methodological focus, the next decade can bring a rich variety and diversity of methods to the e-government landscape.

A diversity of topical content can easily lend itself to a diversity of methodological applications, in particular to mixed methodological strategies. A mixed methodology can align different elements of an e-government project into a single harmonious entity. Synergy occurs when different elements are combined together resulting in a single coherent component [14]. Through a combination of different methodologies, synergy across the various interconnected topics in e-government is possible and can provide deeper insights into the field as a whole. For example, a mixed method approach can combine both an explanation and a description of the diffusion of a specific e-government technical innovation.

Moreover, e-government projects consist of both technical and behavioral components and are essentially socio-technical systems that perform optimally when there is a synergistic fit between the technology and the organization. In e-government, the organizational component can be complex because of underlying political and economic factors that can directly impact a project's staffing, budgeting, development, implementation, and maintenance. Both quantitative and qualitative methods can be used in the same study to better understand the complex and dynamic nature of e-government. One recent e-government study first used statistical techniques to collect data from e-government websites and then used the statistical results as input for two qualitative case studies [7]. A second study used empirical data along with a case study to understand e-government project management [8]. Mixed methodology e-government studies, though currently small in number, highlight the applicability and potentially positive contributions to the field.

Quantitative research can support the empirical data that might be absent in purely qualitative research. Similarly, qualitative research can add more value, meaning, and understanding to a set of numerical findings produced from quantitative studies. Combining quantitative and qualitative methods in the same study is sometimes referred to as "triangulation" and is beneficial for several reasons including expansion of the scope of the study, emergence of new ideas, and complementary perspectives within the study [3]. In e-government, a mixed method study can employ a sequential strategy where either quantitative is followed by qualitative (or vice versa) or a concurrent strategy where both methods are executed simultaneously or one methodology is nested within the other [7].

Within each of the four main categories of e-government (G2C, G2B, G2E, G2G) there are opportunities for mixed method research: qualitative followed by quantitative, quantitative followed by qualitative, qualitative and quantitative simultaneously, qualitative embedded in quantitative, and quantitative embedded in qualitative [7]. Given the goals and objectives of the e-government research project, one mixed method may be more appropriate than the other. The researcher must ultimately determine the level of interaction among the different methods. The field of e-government provides an opportunity for research that is both captivating and challenging. Employing a mixed methodology can be the glue that brings about greater understanding and stimulates growth in the field from both practical and theoretical perspectives.

The mixed methodology perspectives may seem to diverge from existing norms and patterns and can encounter challenges before it can be widely accepted and applied in e-government. As mentioned earlier in the paper there are many advantages of applying mixed methods research in the area of e-government. Of course, there are challenges such as the length of time to complete the study and the variety of resources needed that can deter utilization of a mixed methods approach. Formulating an appropriate research action plan involving mixed methodologies can also be daunting for the researcher who has knowledge and expertise using only a single method. However, these initial challenges will not deter the mavericks who wish to harness the benefits of engagement with mixed method research. The challenge for e-government researchers in the future is to go beyond single methodology studies and focus on framing e-government research encompassing both inductive and deductive methods.

PROPOSED RESEARCH METHODOLOGY

One way to effectively analyze the methodological preferences in e-government research is to conduct a bibliometric study. Bibliometric studies use artifacts such as research notes, conference proceedings, journal articles, and databases to collect and analyze documents on a particular subject [12]. Inter-rater tests are also conducted to confirm consistency of the findings. Some bibliometric studies use specific bibliometric databases and citation databases for data collection. To minimize the inherent challenges of lack of consistency of findings and inclusion / exclusion of specific articles, a broad enough time frame along with multiple data sources should be considered when utilizing a bibliometric research method [6].

For this study a bibliometric approach will be conducted in two phases. Phase one involves a compilation of method that have been used and/or can be used in the area of e-government. Phase two of this project involves the selection of databases, journals, and conferences that publish research in the area of e-government. The time frame that this project focuses on is the period 1999-2009. This represents the recent past of e-government. Past studies will be analyzed for methodological variability and also used to identify trends that have occurred that can be used as a guideline for areas of future growth in this emerging field. We expect that there will be a limited number of mixed method research in the past literature largely because of the inherent challenges such as cost and time that are associated with such studies.

CONCLUSION

The field of e-government is new and diverse. There are many technical, social, economic, organizational, and political factors that impact the development of e-government projects. From the rich tradition of diverse fields, diverse theories and diverse research streams, it would seem myopic to be restricted to specific philosophical views in the area of e-government. Researchers in e-government would be remiss in not embracing the potential value of pursuing mixed method studies. E-government practitioners such as developers and analyst can also benefit by exploring both qualitative and quantitative methods concurrently, as they build and implement new e-government initiatives. Combining quantitative and qualitative data collection methods, can also serve as part of post-implementation assessment and evaluation of e-government projects. When effectively applied, a mixed method approach can provide more meaningful insights into the rich complexity underlying e-government. The future of e-government research can potentially be significantly enhanced by continued and increased diversity of research methods.

REFERENCES

[1] Andersen, K. V., and H. Z. Henriksen. "The First Leg of E-government Research: Domains and Application Areas 1998-2003," *International Journal of Electronic Government Research*, 2005, 1(14), 26-44.

[2] Cao, J..J. M. Crews.M. Lin.A. Deokar.J. Burgoon, and J. F. Nunamaker Jr. "Interactions Between System Evaluation and Theory Testing: A Demonstration of the Power of a Multifaceted Approach to Information Systems Research," *Journal of Management Information Systems,* 2006, 22(4), 207-235.

[3] Creswell, J. W., and V. L. Plano Clark. 2007. Designing and Conducting Mixed Methods Research. Sage, Thousand Oaks, California.

[4] Dwivedi, Y. K. "An analysis of e-Government research published in Transforming Government: People, Process and Policy (TGPPP?)," *Transforming Government : People, Process and Policy,* 2009, 3(1), 7-15.

[5] Evans, D., and D. C. Yen. "E-Government: Evolving relationship of citizens and government, domestic, and international development," *Government Information Quarterly,* 2006, 23(2), 207-235.

[6] Frandsen, T. F., and J. Nicolaisen. "Intradisciplinary Differences in Database Coverage and the Consequences for Bibliometric Research," *Journal of the American Society for Information Science and Technology,* 2008, 59(10), 1570–1581.

[7] Gil-Garcia, J. R., and T. A. Pardo. 2006. Presented at the Proceedings of the 39th Annual Hawaii International Conference on System Sciences Hawaii.

[8] Ke, W., and K. K. Wei. "Understanding E-government Project Management: A Positivist Case Study of Singapore," *Journal of Global Information Technology Management,* 2006, 9(2), 45-61.

[9] Lee, S. M..X. Tan, and S. Trimi. "Current Practices of Leading E-government Countries," *Communications of the ACM*, 2005, 48(10), 99-104.

[10] Meneklis, V., and C. Douligeris. "Bridging theory and practice in e-government: A set of guidelines for architectural design," *Government Information Quarterly,* 2010, 27(70-81.

[11] Miles, M. B., and M. A. Huberman. 1994. Qualitative Data Analysis: An Expanded Sourcebook, Second ed. Sage.

[12] Schneider, J. W., and P. Borlund. "Introduction to bibliometrics for construction and maintenance of thesauri: Methodical considerations," *Journal of Documentation,* 2004, 60(5), 524-549.

[13] Sekaran, U. 2000. Research Methods for Business: A Skill-Building Approach. John Wiley & Sons, New York.

[14] Senge, P. 2006. The fifth discipline : the art and practice of the learning organization. Doubleday/Currency, New York.

[15] Siau, K., and Y. Long. "Synthesizing E-government Stage Models: A Meta-synthesis based on Meta-ethnography Approach," *Industrial Management & Data Systems,* 2005, 105(3/4), 443-458.

[16] West, D. M. "E-Government and the Transformation of Service Delivery and Citizen Attitudes," *Public Administration Review,* 2004, 64(1), 15-27.

IT ISSUES OF CITIBANK - A CASE STUDY

Gregory Amato, Richard Stockton College of New Jersey, gamato022@hotmail.com
Haidong Liu, Richard Stockton College of New Jersey, haidong.liu@stockton.edu

On June 16, 1812 with two million in capital, City Bank of New York (now Citibank) opened for business in New York City. In 1998, all Citi divisions merged with Travelers Group to form the colossal financial supermarket today, known as Citigroup. Today Citibank stands as the largest bank in the world operating in over one hundred countries. The financial services company was built to create a highly diversified financial services company that could act as a one stop shop for client solutions around the world.

As a commercial relationship manager, Mike Adam's responsibility is to acquire new clients to do business with Citibank. In order to persuade someone to move their existing banking relationship to another institution it is important to convey how Citibank stands out from all other banks. Mike is often asked, "What makes your bank different from a bank such as Wachovia?" The answer is simple, "We do business in over 100 countries and allow you to perform all of your financial services at one institution, and no other financial institution can offer that." After working for 2 divisions at Citigroup (Citicapital & Citibank) Mike has learned that Citi's biggest strength is also its biggest weakness and the obstacles we face on a daily basis stem mostly from technological issues.

Citibank's Technology Issues

The information technology challenge Mike face is that the business has 13 different subsidiaries within Citigroup's financial services firm. When a customer walks into the bank and wants to check their credit card balance Mike has to refer them to the credit card affiliate, Citicards. If customers want to view their mortgage statement they are referred to the mortgage affiliate, Citimortgage. There are eleven other sub-entities that Mike can refer any of valued customers. This is not exactly one-stop shopping. Citi's other huge strength is their international presence. Citibank offers bank locations throughout every major city in the world. But what happens when someone with a Citibank Asia account wants to enter a Citibank North America branch? The international customer is unable to make deposits into their foreign account. The bank doesn't have a universal banking system that allows you to make deposits into its overseas Citibank account.

Implementing a software solution that can give capability to view accounts across businesses and countries would be an IT solution that would truly make the stated benefits Citibank offers to its customers more adequate and unique.

Solution - ERP

It is an overwhelming challenge for a business to excel in every enterprise, especially a multi business international company like Citigroup. An enterprise resource planning (ERP) system provides a foundation for collaboration between departments, enabling people in different business areas to communicate. ERP integrates all departments and functions across a company onto a single computer system that can serve all those different departments' particular needs.

The heart of an ERP system is a database that collects information and feeds it into all the ERP system's individual applications. When a user enters or updates information in one module, it is immediately and automatically updated throughout the entire system allowing everyone within the company to view the same information across the database. Citigroup has traditionally isolated information from within different subsidiary businesses as well as different countries. Citibank North American operates on a separate database than Citibank Europe, and Citicards operates with a different database than Citimortgage. ERP is the logical solution to the disorder of incompatible applications in the different businesses. It addresses the need for global information sharing and reporting, and is used to avoid the pain and expense of fixing legacy systems.

Having everyone involved in sourcing, producing, and delivering the company product works with the same information. This eliminates redundancies, reduces wasted time, reduces manpower, and removes misinformation. Eliminating and reducing these will reduce costs, streamline customer service, and keep you competitive in the financial services industry.

Not only is this beneficial to your bottom line and your customer service but ERP systems can greatly improve your internal operations. The three common core ERP components are: Accounting and finance, production and materials management, and human resources.

Accounting and finance ERP components manage accounting data and financial processes within the enterprise with functions such as general ledger, accounts payable, accounts receivable, budgeting, and asset management. Reforming the accounting and finance within the segregated businesses operating within Citi can give you a more real-time view of your revenues, expenses, profits and losses. Having a more timely feedback can really provide an edge in planning when to cut expenses or make necessary changes to protect your bottom line. Production and materials management ERP components handle the production planning and execution such as demand forecasting, production scheduling, cost accounting and quality control. Citigroup has 330,000 employees. Human resources ERP components track employee information such as payroll, benefits, compensation, and performance assessment. This component can identify which employees are utilizing the appropriate systems to have a greater impact on profit.

It is important to weigh your pros with your cons when deciding on implementing a new IT system. Citibank will have to first purchase the software, and then hire consultants to install the software accordingly. Once the software is installed it must be customized to meet the organizations needs. After the setup you have to transfer the old data to the ERP system and train your employees on operational procedures. A lot of time and money is spent to implement such a system. On average the company's investment into the system will not show any benefits for 8-18 months down the road but the cost savings for implementing new ERP systems are $1.6 million on average.

Business Innovation – What's new in banking

The core issues faced by banks today are on the fronts of customer's service expectations, cutting operational costs, and managing competition. Technology can help banks in meeting these objectives. IT is central to banking. It has moved from being just a business enabler to being a business driver. In the banking and financial services sector, being the early adopters of any new technology, defines the roadmap for future technology adoption.

Banks are focused on three areas: meeting customer's service expectations, cutting costs, and managing competition. For this banks are exploring new financial products and service options that will help them grow without losing existing customers. And any new financial product or service that a bank offers will be intrinsically related to technology.

Remote Cash Capture

Another way to remain competitive for Citibank is to be ahead of the technology curve among your competitive peers. Remote cash capture (RCC) is an emerging technological feature being piloted by larger commercial banks. Just as banks are beginning to offer check imaging, this technology offers a similar electronic deposit service for cash on site that electronically deposits funds. Its primary benefit is to streamline the traditional cash handling processes. RCC eliminates business cash offices, improves cash availability and liquidity and reduce labor, theft and shortages. It uses cash cycle control, electronic information exchange, and data management strategies to eliminate cash handing associated with store safes. It simplifies cash management, POS reconciliation, and bank depositing.

Remote cash capture works by supplying retailers with physical cash recycling systems with electronic connections to their financial institutions. Cash is handled to and from the centralized RCC container. As funds are distributed and deposited the associated data is transmitted daily to their banks. The bank periodically collects cash from the remote capture systems just as they would at an ATM. This is beneficial to businesses to reduce theft and shortages, manage funds availability, and eliminate deposit discrepancies. As more banks offer remote cash capture businesses will benefit by bringing the bank to the local business. It removes bank barriers and allows retailers to focus banking relationships on service and costs no matter how far they are from their financial institution.

Mobile Banking

It is important to understand that mobile banking is not the same as online banking. The pace of growth has far exceeded online banking to where mobile banking has become do-or-die for banks. The immediate access and control over one's finances has been the preferred channel in today's economy. People feel more secure banking from their mobile phones than on their PC's because of all the viruses and malware that can infect computers. The mobile phone is also location sensitive providing an additional security advantage over online banking. Phones such as the iphone provide not only security but a marketing opportunity through the iphone's application store. There are nearly two billion mobile phones in the world today. Both the wireless and financial industries would benefit from enabling consumers to use their cell phones to do their banking. For wireless carriers, it would represent another way to potentially garner revenue. Meanwhile, banks that choose to participate can offer a competitive advantage through the convenience of mobile banking. Citibank should really capitalize on the growth from mobile banking given their global presence and the number of mobile phones worldwide.

Check Scanning ATM's

Bank customers have long been weary from depositing an envelope of cash or checks into an ATM machine. Bank of America has led the recent innovation of ATM technology.

They have implementing a system that provides image scanning of cash and checks as they are inserted into the machine. This system provides security and peace of mind to the user through confirmation of the amount of your cash deposit, or a printout of any checks that you deposit on your ATM receipt. Citibank has begun piloting the new ATM's at Citibank branches in California and is hoping to roll out the new ATM's in Citibank branches and 7-11 stores by the end of 2009. Besides security another benefit here will come in the form the increased efficiency at Citibank. Not only will this reduce paper shuffling and bank errors, but it can reduce expenses at the branch level. If you have the ease of depositing items at an ATM, the bank's hours can be cut and there will be less staff needed per transaction. Citibank has an advantage over Bank of America in this technology segment because of their Citibank ATM's located at every 7-11 store in the country. 7-11 stores are open 24 hours a day and having the ability to make deposits at your local corner 7-11, twenty four hours a day is yet another distinct competitive technology advantage over rival banks.

The U.S. banking sector is facing its toughest times in American history. Every dollar being spent is scrutinized causing little investment into the future. Unfortunately through these hard hit times financial institutions are losing their most talented individuals to competitors. Banks like Citibank should invest now in performance management systems to identify internal and external talent to be in place for market recovery. ICIMS is a talent platform solution which supports more than 700 corporations worldwide. ICIMS manages applicant tracking capabilities, on boarding and off-boarding, performance management, succession planning, workforce planning and more via a single web based system. Using ICIMS eliminates the need to purchase costly talent management solutions with multiple software that may not compatible with their current systems. ICIMS provides analysis to help better identify the company's core objectives and then tie in measurable results of how individuals are meeting these objectives.

Automation is the basic thing that banks need to have in place. It involves a combination of centralized networks, operation, and core banking applications. As mentioned earlier, IT is crucial to banking. It has moved from being just a business enabler to being a business driver. In the banking and financial services sector, Citi must take the appropriate steps to eliminate inefficiencies in back-end operations. The company must consolidate its IT systems across businesses and be innovators with 21st century technology. The term "Digital Darwinism" implies that organizations that cannot adapt to the new demands placed on them for surviving in the information age are doomed to extinction. Being the early adopters of any new technology, defines the roadmap for future technology adoption. Citigroup has been a dominant financial services company since the 1800's. The recent turmoil in the financial meltdown has temporarily crippled the business. Through restructure, sale of non profitable assets, expense reduction, technology, and the proper investments Citigroup is soon to emerge back as a leaner and more profitable franchise.

REFERENCES

[1] Baltzan, Paige, Phillips, Amy, Haag, Stephen. "Business Driven Technology". McGraw-Hill Irwin 2008.

[2] www.tech-fag.com/-,glyrprise-resource-planning-erp.shtml

[3] www.americanbanker.com/emailthis.html

[4]http://www.remotedepositcapture.com/new/news.aspx?aid=2550

[5] http://online.wsj.com/article/SB123362222507641723.htrn

Panel Session for NEDSI

Title: NEDSI: Why Should You Be Involved?

Panelists: Paul Mangiameli, University of Rhode Island
Dan Reid, University of New Hampshire
Laura Forker, University of Massachusetts-Dartmouth
Eric Stein, Penn State University-Great Valley

This interactive session discusses some of the different ways in which members can become more involved in NEDSI. Activities include reviewing manuscripts, being a session chair or track chair, a local arrangements coordinator, a member of the board of directors, Proceedings Editor, VP of Promotional Activities, etc. The advantages and disadvantages of such involvement will be discussed as well as the general progression of responsibility within NEDSI. The primary purpose of the session is to encourage and recruit newer members for service to NEDSI. The panel consists of two past-Presidents of NEDSI, along with two newer members of the NEDSI Board of Directors.

ZIPF DISTRIBUTION OF INNOVATION: FIRMS AND CITIES

Thomas F. Brantle, Stevens Institute of Technology, (732) 872-7742,
thomas.brantle@stevens.edu

ABSTRACT

The size distribution of firms and cities globally is investigated using the number and size of a firm's and city's constituent 'innovation' components. Innovation as characterized by patents, citations and inventor activity is used as a measure of firm and city size. Patents, citations, and inventor activity have been considered excellent proxies for the innovation process, technological advancement and resultant economic growth. It is shown that the size distribution of firms and cities globally is well represented by a Zipf distribution with exponent equal to approximately "1" across several size measures of innovation production and activity, including a wide range of firm and city sizes, from the smallest to the largest. These results are confirmed empirically. A direct connection is established between the structure of organizational and geographic growth as represented by innovation production and activity.

Keywords: Zipf distribution, innovation, patents, firms, cities

INTRODUCTION

The size distribution of firms and cities in industrial countries has been shown to be highly skewed to the right, such that small numbers of large firms and cities coexist alongside larger numbers of smaller firms and cities. This observed skewness has been stable over time: being insensitive to changes in political and economic environments; and independent of mergers, acquisitions or entries and bankruptcies of firms, or mobility their employees; as well as demographic changes in city populations, or birth and death rates and mobility of its citizens. [2] [4] [5] [6] [7] [8] [9] [12] [14] [15] [16] [17] [18] [21] [22] [25] [30] [31] [32] [34]

Earlier research has hypothesized and demonstrated that the agglomeration and residential mobility of the population between different geographic locations is tightly connected to associated economic activity. Thus, a full understanding of organizational and geographic activity involvers understanding population mobility along with associated economic driving forces (e.g., innovation and technology). The firm and city size distributions are relevant indicators of industrial and population concentrations, respectively, quantities of interest to both technology and financial managers as well as antitrust and public policy decision makers. [4] [7] [13] [20] [21] [22] [23] [26] [28] [29]

Firm size has been typically measured by total assets, sales revenue, market value or number of employees and city size determined by population census or income figures. Previous studies have differed as to whether the Lognormal distribution or the Zipf (or Pareto) power law distribution (a.k.a., the Rank-Size rule) provides an appropriate model, as well as to their consideration of only large firms and cities versus the inclusion or exclusion of increasing numbers of medium and smaller sized firms and cities in their investigations. [2] [4] [5] [6] [7] [8] [9] [12] [14] [15] [16] [17] [18] [21] [22] [25] [30] [31] [32] [34]

The Zipf or Pareto power law distribution or scaling distribution is defined as the complimentary cumulative distribution (CCDF) where:

$$P(S > s) \sim s^{-\alpha}, s > 0 \; and \; \alpha > 0$$

The special case of $\alpha = 1$ is known as the Zipf distribution [34]. The Pareto and Zipf distribution or the rank-size distribution has described notable regularity in many phenomena including the distribution of incomes and wealth among individuals, particle sizes of sand, meteor impacts on the moon, lengths of rivers, frequencies of word usage in addition to firm and city sizes. The Zipf or Pareto power law distributions are typically estimated by first constructing the rank size distribution by ranking the raw data values from largest to smallest with the fractional rank distribution providing the CCDF directly. Ordinary Least Squares (OLS) or Maximum Likelihood Estimation (MLE) is typically applied to estimate the scaling exponent α. [2] [10] [24] [27] [34]

The association between the distributions of economic activity, as represented by technological innovation, is investigated in this paper across geographic locations (e.g., cities) and organizations (e.g., firms). Specifically, the global size distribution of firms and cities as represented by innovation production and activity is considered in this research. Innovation has long been seen as driving productivity and economic growth. Patents have long been explored and recognized as a very useful and productive source of data for the assessment of innovation development and technological advancement.

INNOVATION AND PATENT DATA

Patents provide a wealth of information and a long time-series of data about inventions, inventors, collaboration, prior knowledge, and the assigned owners and locations (e.g., firms and cities). Patents and the inventions they represent have several advantages as a technology indicator. In particular, patents and patent citations have long been recognized, and more recently invention collaboration, as a very rich and fertile source of data for studying knowledge, innovation and technological change. As such, they provide a valuable tool for public and corporate technology analysis, planning and policy decisions. [3] [11] [19]

Patents and patent citations constitute a documentation path for knowledge creation, technology transfer and innovation flows. When one patent cites another patent the citation indicates knowledge and information flowing from the cited patent to the citing patent. Therefore, the knowledge and innovation information made publicly available by the cited patent has not only flowed to the citing patent but has also influenced both the citing patent's inventor(s) and the invention it represents. Two sets of measures are generally considered: forward-based measures that are derived from the relationship between a given patent and subsequent technology developments that build upon it, which are considered descendents; and backward-based measures that are derived from the relationship between a given patent and the body of knowledge that preceded it, which are considered antecedents. The supposition is that forward-based measures would be informative of the subsequent impact of research outcomes, whereas backward-based measures would be informative of the nature of the research. Forward citations or citations received measure originality and are an indication that the technology area is new or less-developed and the innovation is more likely to be significant and of greater value. Backward citations or citations made measure generality and are an indication that the technology area is well-developed and the innovation is more likely to be derivative and of less value. [19] [33]

The United States Patent and Trademark Office (USPTO) utility patent, citation and inventor raw data files are used in this study. [1] First, all the data is appropriately cleaned, filtered and reformatted for data analysis purposes. The data comprises detailed information on approximately 3.1 million U.S. patents granted between January 1975 and December 2005. Each patent contains highly detailed information on the invention itself, the technological area to which it belongs, the inventors, including their geographic location, and a company or organizational identifier. Additionally it includes the list of all citations made by patents issued between 1975 and 2005; about 21.6 million citations that either cites are cited by other patents. In addition the patent inventor data files contain approximately 6.6 million unique individual patent-inventor pairings, consisting of about 3.1 million patents and roughly 1.9 million individual inventors, from 1975 through 2005.

For each patent the geographic location (country, state and city) is determined by the "first" or "primary" inventor on the patent. The company or organizational identifier of each patent is known and recorded for approximately 84% of the set of patents, with the remaining 16% generally referred to as the "independent inventors" which includes small businesses, self-employed inventors and other small organizations. All these patents and inventors represent companies, organizations and individuals across both private industry and government organizations, with about 55% of patents produced and owned by U.S. entities and the remaining 45% spread across the globe and numerous countries.

To account for the patents associated with the "independent inventors" an additional organizational identifier is constructed. For each of these "independent inventors" patents its "first" inventor is considered equivalent to a small business owner or organizational head and assigned its own unique organizational identifier. Thus all patents produced under this individual are considered owned by this small business or organization. Similarly, all inventors associated with these independent inventor patents are considered employees or associates of this small business or organization. Thus, two organizational identifiers are identified and/or constructed, the first is the original assignee heretofore denoted as "Companies" and the second which includes the known original assignees plus the independent inventors will be referred to as "Firms". Hence Firms includes the Companies as well as the independent inventor small businesses, self-employed, as well as other small organizations.

In addition, two versions of Firms and Companies are considered, the first is unrestricted by geographic boundaries and these are referred to as Global Firms and Global Companies, respectively. The second restricts Firms and/or Companies to either their State boundaries for U.S. owned patents and or its Country boundaries for non-U.S. owned patents, and these are referred to as Regional Firms and Regional Companies. Thus, the Regional Firm and Company identifier represents the local or regional presence of an organization while the Global Firm and Company identifier represents its worldwide or global reach.

Cities are identified and recorded for each Patent based on the city of the “first” or “primary” inventor and is considered the geographic location of the invention. While, individual inventor’s Cities are identified and recorded based upon the location and city of residence of the inventor at the time the patent was granted. Thus individual inventors may be associated with different Cities over time and their patenting production. Similarly, inventors may be associated with different Global Firms and Companies or Regional Firms and Companies over time and their patenting production.

ZIPF DISTRIBUTION OF INNOVATION – RESULTS DISCUSSION

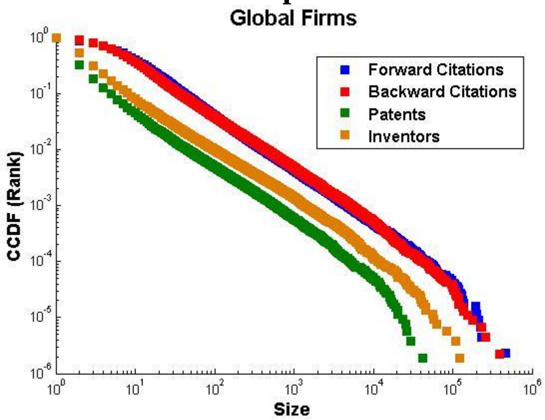
The firm and city size distribution is considered over a 31-year period of innovation, from 1975 to 2005, for several size measures of innovation. Firm and city innovation activity is measured by patenting and inventor production and activity, specifically the number of forward citations (cited patents, an output measure of originality, productivity and value), backward citations (citing patents, an input measure of generality, derivation and cost), total number of patents (an output measure of productivity and value), number of inventor instances associated with patent production (an input measure of human capital). Two definitions of a firm’s boundaries are considered, one which represents its global innovation activity and another restricting it to its country or state boundaries. Importantly all sizes of firms and cities are considered, including the full recognition of independent inventors (e.g., small businesses and individuals) which represents roughly one-sixth of patents and invention activity.

Figures 1, 2 and 3 respectively provide the CCDF or rank-size distributions plots across Global Firms and Companies, Regional Firms and Companies, and for All Cities, for several measures of innovation activity, forward citations, backward citations, total number of patents, and total number of inventors (instances) on double-logarithmic plots. Visually it is quite apparent that a straight-line provides an excellent fit to the Global as well as the Regional Firms and Companies plots and thus an appropriate Zipf power-law fit, with each measure of innovation providing similar and parallel results. For the All Cities curves there is a slight curvature and thus a power-law with exponential cutoff or perhaps a stretched exponential may provide an alternative model and fit. However, since most studies of the size distribution of cities have typically measured MSAs (Metropolitan Statistical Areas) or other larger local geographic boundaries (e.g., counties) rather than simply the city proper as utilized here it is conjectured that this increased aggregation would provide an even better Zipf power law fit than the that observed here.

Table 1 summarizes the results of OLS estimation for the Zipf power law exponents across, Global and Regional Firms and Companies plus All Cities for all four measures of innovation size considered. Almost all the estimates are within +/- 0.05 of a Zipf exponent equal to 1, showing strong agreement with previous studies of the Zipf distribution of firms, companies and/or cities. MLE estimation utilizing appropriate lower and/or upper cutoff points provide similar results.

FIGURE 1

Global Firms and Companies



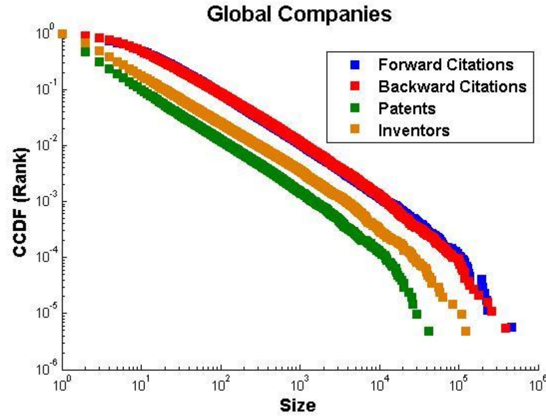
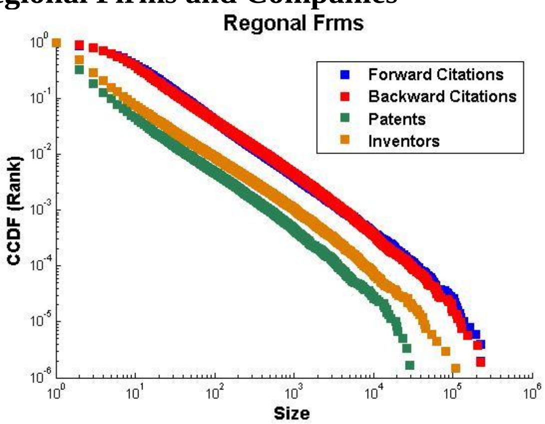


FIGURE 2

Regional Firms and Companies



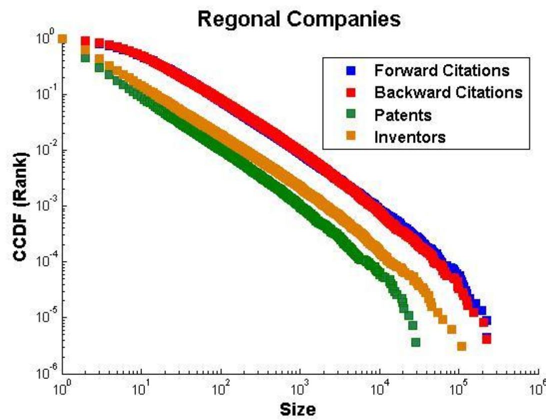


FIGURE 3

All Cities

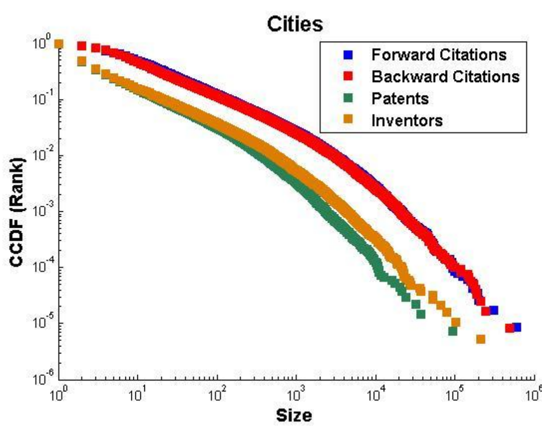


TABLE 1

OLS Zipf Rank-Size CCDF Exponent Estimates

OLS Zipf Rank-Size CCDF Exponent Estimates					
Innovation Size Measure	**Global Firms**	**Global Companies**	**Regional Firms**	**Regional Companies**	**All Cities**
Forward Citations	0.9677	0.9275	1.0192	0.9850	0.8896
Backward Citations	0.9678	0.9406	1.0393	1.0160	0.8881
Total Patents	0.9958	0.9826	1.0756	1.0641	1.0660
Total Inventors	0.9472	0.9374	1.0163	1.0085	0.9879

SUMMARY AND CONCLUSIONS

It is generally found that in all cases, across innovation size measures, that the empirical firm and city size distribution for innovation production and activity is well represented by a Zipf power law function with exponent equal to approximately "1" based on both ordinary least squares (OLE) and maximum likelihood estimation (MLE) estimation of the rank size distribution. Thus evidence is provided of scaling behavior for both the firm and city size distribution independent of the definition of innovation size chosen. The existence of the Zipf power law is also seen to be equally robust for firms and cities across measures of innovation and size. As a result, firm and city size are both inversely proportional to the rank of the size of firm and city, respectively, regarding economic activity as represented by technological innovation production and activity. An interesting property of the firm and city size distribution observed is that independent of how "innovation size" is defined the distributions remain similar, although the position of individual firms and cities does depend on the definition of innovation size, while the shape of the size distribution does not.

This research represents a unique and first time contribution to the economic, organization, innovation, and firm and city size distribution literature as it represents firms and cities globally, includes not only the largest firms and cities but also the smallest firms and cities in terms of innovation production and activity, and importantly is the first time the size distribution for firms and cities has been so described and analyzed for innovation production and activity, i.e., patenting and inventor activity and production. It also indicates that the size distribution of innovation production and activity for firms and cities, i.e., across geography and organization, is robust and follows a parallel pattern and structure. Thus, a direct connection is established between the size structure of organizational and geographic growth as represented by innovation production and activity.

REFERENCES

[1] *United States Patent and Trademark Office (USPTO) Technology Assessment and Forecast (TAF) Database*. 2005, United States Patent and Trademark Office (USPTO), Electronic Information Products Division: Alexandria, VA 22313.

[2] Axtell, R.L., *Zipf Distribution of U.S. Firm Sizes.* Science, 2001. **293**: p. 1818-1820.

[3] Cohen, W.M. and S.A. Merrill, eds. *Patents in the Knowledge-Based Economy*. 2003, The National Academic Press: Washington, DC.

[4] Cordoba, J.-C., *On the Distribution of City Sizes.* Journal of Urban Economics, 2008. **63**: p. 177-197.

[5] Dobkins, L.H. and Y.M. Ioannides, *Evolution of the U.S. City Size Distribution*, in *The Economics of Cities*, J.-M. Huriot and J.-F. Thisse, Editors. 1998, Cambridge University Press: New York, NY.

[6] Eeckhout, J., *Gibrat's Law for All Cities.* American Economic Review, 2004. **94**(5): p. 1429-1451.

[7] Fujita, M., P. Krugman, and A.J. Venables, *The Spatial Economy:Cities, Regions and International Trade*. 2001, Cambridge, MA: The MIT Press.

[8] Gabiax, X., *New Ideas on Economic Growth: Zipf's law and the Growth of Cities.* The American Economic Review, 1999. **89**(2).

[9] Gabiax, X., *Zipf's Law for Cities: An Explanation.* The Quaterly Journal of Economics, 1999. **114**(3): p. 739-767.

[10] Gabiax, X. and R. Ibragimov, *Rank - 1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents*, in *NBER Working Paper*. 2009.

[11] Griliches, Z., *Patent Statistics as Economic Indicators: A Survey.* Journal of Economic Literature, 1990. **28**(4): p. 1661-1707.

[12] Hart, P.E. and S.J. Prais, *The Analysis of Business Concentration: A Statistical Approach.* Journal of the Royal Statistical Society, 1956. **119**: p. 150-191.

[13] Henderson, J.V., *Urbanization and Growth*, in *Handbook of Economic Growth*, P. Aghion and S.N. Durlauf, Editors. 2005, Elsevier B.V.: Amsterdam.

[14] Ijiri, Y. and H.A. Simon, *Business Firm Growth and Size.* American Economic Review, 1964. **54**(2): p. 77-89.

[15] Ijiri, Y. and H.A. Simon, *Effects of Mergers and Acquisitions on Business Firm Concentration.* Journal of Political Economy, 1971. **79**(2): p. 314-322.

[16] Ijiri, Y. and H.A. Simon, *Interpretations of Departures from the Pareto Curve Firm Size Distributions.* Journal of Political Economy, 1974. **82**(2): p. 314-322.

[17] Ijiri, Y. and H.A. Simon, *Skew Distributions and the Sizes of Business Firms*. 1977, New York: North Holland.

[18] Ioannides, Y.M. and H.G. Overman, *Zipf's Law for Cities: An Empirical Examination.* Regional Science and Urban Economics, 2003. **33**: p. 127-137.

[19] Jaffe, A.B. and M. Trajtenberg, eds. *Patents, Citations, and Innovations: A Window on the Knowledge Economy*. 2002, MIT Press: Cambridge, MA.

[20] Krugman, P., *Development, Geography and Economic Theory*. 1995, Cambridge, MA: MIT Press.

[21] Krugman, P., *Confronting the Mystery of Urban Hierarchy.* Journal of the Japanese and International Economies, 1996. **10**: p. 399-418.

[22] Krugman, P., *The Self Organizing Economy*. 1996, Malden, MA: Blackwell Publishers Ltd.
[23] Lucas, R.E.J. and E. Rossi-Hansberg, *The Internal Structure of Cities.* Econometrica, 2002. **70**(4): p. 1445-1476.
[24] Newman, M.E.J., *Power Laws, Pareto Distributions and Zipf Law*. Contemporary Physics, 2005. **46**(5): p. 323-351.
[25] Overman, H.G. and Y.M. Ioannides, *The Cross Sectional Evolution of the U.S. City Size Distribution.* Journal of Urban Economics, 2001.
[26] Porter, M.E., *Clusters and the New Economics of Competition.* Harvard Business Review, 1998. **78**: p. 77-90.
[27] Reed, W.J., *The Pareto, Zipf and Other Power Laws.* Economics Letters, 2001. **74**: p. 15-19.
[28] Rosenthal, S.S. and W.C. Strange, *Geography, Industrial Organization, and Agglomeration.* The Review of Economics and Statistics, 2003. **85**(2): p. 377-393.
[29] Saxenian, A., *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. 1994, Cambridge, MA: Harvard University Press.
[30] Simon, H.A. and C.P. Bonini, *The Size Distribution of Business Firms.* American Economic Review, 1958. **48**(4): p. 607-617.
[31] Stanley, M.H.R., et al., *Zipf Plots and the Size Distribution of Firms.* Economics Letters, 1995. **49**: p. 453-457.
[32] Steindl, J., *Random Processes and the Growth of Firms*. 1965, London: Griffin.
[33] Trajtenberg, M., R. Henderson, and A.B. Jaffe, *University versus Corporate Patents: A Window on the Basicness of Innovation.* Economics of Innovation and New Technology, 1997. **5**: p. 19-50.
[34] Zipf, G.K., *Human Behavior and the Principle of Least Effort*. 1949, Reading, MA: Addison Wesley.

ASSESSING GRADIENT BOOSTING IN THE REDUCTION OF MISCLASSIFICATION ERROR IN THE PREDICTION OF SUCCESS OF ACTUARIAL MAJORS

Alan Olinsky, PhD – Bryant University, Smithfield, RI, 401-232-6266, aolinsky@bryant.edu
Kristin Kennedy, PhD – Bryant University, Smithfield, RI, 401-232-6316, kkennedy@bryant.edu
Bonnie Brayton Kennedy, CPA, MS - Salve Regina University, Newport, RI, 401-295-7800, bonnie.kennedy@salve.edu

ABSTRACT

This paper provides a relatively new technique for predicting the retention of students in an actuarial mathematics program. The authors utilize data from a previous research study. In that study, logistic regression, classification trees, and neural networks were compared. The neural networks and classification trees were most accurate and were able to deal with missing data quite successfully. In this paper, we examine the use of gradient boosting to improve the accuracy of classification trees. We focus on trees as opposed to neural networks as the results are transparent and easily understood.
Key Words: Logistic Regression, Data Mining, Neural Nets, Decision Trees, Gradient Boosting

BACKGROUND

This paper provides an analysis of data pertaining to the retention of students in an actuarial mathematics program. In a previous paper [4], data mining techniques were applied to a study that investigated the likelihood that incoming college freshmen majoring in actuarial mathematics (AM) actually graduate in this major. In completing this study, the authors obtained data from 328 Bryant University graduating students who had begun their college careers as freshman actuarial majors. A logistic regression, decision tree, and neural network were used to predict the successful completion of the program based on SAT scores, high school rank in class and a college mathematics test. There were a fair number of missing values in this data set. The authors re-analyze data from this study.

INTRODUCTION

In the present study, the same data is examined using gradient boosting in an attempt to further reduce the misclassification error obtained with decision trees. Variables that were available and used as predictors included Math SAT, verbal SAT, percentile rank in the students' high school graduating class, and his/her percentage score on a college mathematics placement exam.

In this study, we focus on decision trees as the results are transparent and clearly understood (compared to, for example, neural networks). In addition, decision trees can handle missing values whereas logistic regression can only be performed when cases with missing values are excluded from the analysis or imputation occurs first.

As mentioned, decision trees were one of the models used in the prior study. We wanted to see whether the technique of gradient boosting, which was developed by Jerome Friedman [2][3], might improve the fit of the model. Specifically, we look at misclassification error.

Decision Tree Algorithms

Traditional decision trees use algorithms that search for an optimal partition of the data defined in terms of the values of a single target variable. The optimality criterion depends on how this target variable is distributed into the partition segments. The more similar the target values are within the segments, the greater the worth of the partition. Most partitioning algorithms further partition each segment by recursive partitioning. The partitions are then combined to create a predictive model, which is evaluated by goodness-of-fit statistics defined in terms of the target variable [1].

Gradient Boosting

As stated in Enterprise Miner's overview of the Gradient Boosting Node: "Gradient boosting is a boosting approach that resamples the data set several times to generate results that form a weighted average of the resampled data set. Tree boosting creates a series of decision trees which together form a single predictive model. A tree in the series is fit to the residual of the prediction from the earlier trees in the series. This residual is defined in terms of the derivative of a loss function. For squared error loss with an interval target the residual is simply the target value minus the predicted value. Each time the data is used to grow a tree and the accuracy of the tree is computed. The successive samples are adjusted to accommodate previously computed inaccuracies. Because each successive sample is weighted according to the classification accuracy of previous models, this approach is sometimes called stochastic gradient boosting. Boosting is defined for binary, nominal, and interval targets." [1]

It further states: "Like decision trees, boosting makes no assumptions about the distribution of the data. For an interval input, the model only depends on the ranks of the values. For an interval target, the influence of an extreme value theory depends on the loss function. The Gradient Boosting node offers a Huber M-estimate loss which reduces the influence of extreme target values. Boosting is less prone to overfit the data than a single decision tree, and if a decision tree fits the data fairly well, then boosting often improves the fit." [1]

Software

We conducted our analysis with SAS© Enterprise Miner 6.1 [5]. This software can be used as a standalone package on an individual machine or through OnDemand, in which the user logs on to a cloud computer at SAS and runs the software on the server (http://www.sas.com/success/bryantuniversity.html). We have found SAS to be very reasonable in its pricing to academia and they provide their trainers' kits free of charge. Enterprise Miner includes modeling nodes for the decision trees and gradient boosting needed in this study.

Model

This model includes the nodes from the previous study. Hence, the modeling nodes for logistic regression, neural networks, and decision trees appear in Figure 1. It now also includes the node for gradient boosting. The results from this method are compared to those from the original methods.

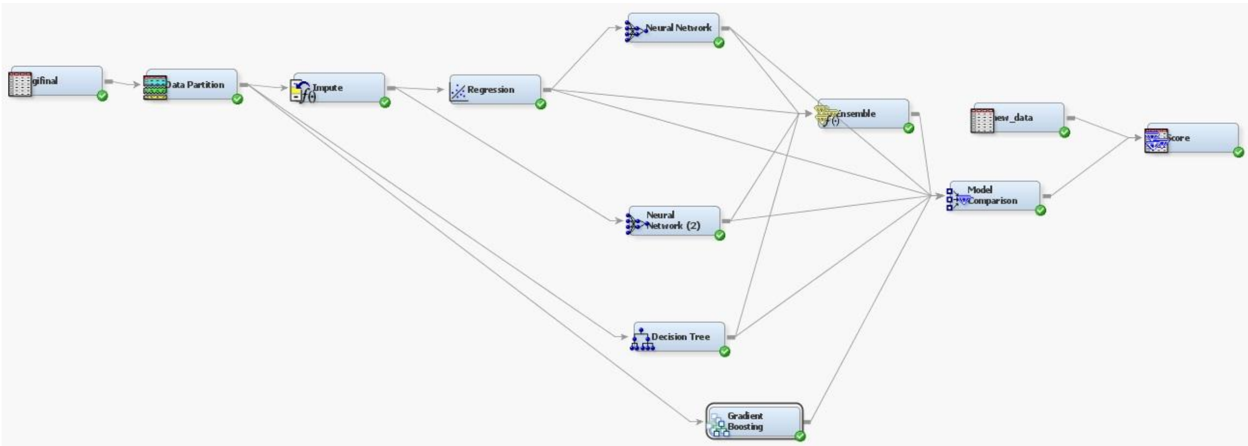


Figure 1. Complete Model in Enterprise Miner

RESULTS

To determine if the technique of gradient boosting improves the results obtained from the decision tree, we first examine the results from the decision tree. Table 1 includes fit statistics, Table 2 has variable importance and Figure 2 displays the actual resulting tree.

Target	Fit Statistics	Statistics Label	Train	Validation	Test
MAJOR	_NOBS_	Sum of Fre...	247	81	.
MAJOR	_SUMW_	Sum of Cas...	494	162	.
MAJOR	_MISC_	Misclassific...	0.299595	0.358025	.
MAJOR	_MAX_	Maximum A...	0.75	0.75	.
MAJOR	_SSE_	Sum of Squ...	101.9684	36.07701	.
MAJOR	_ASE_	Average Sq...	0.206414	0.222698	.
MAJOR	_RASE_	Root Avera...	0.454328	0.471908	.
MAJOR	_DIV_	Divisor for A...	494	162	.
MAJOR	_DFT_	Total Degre...	247	.	.

Table 1: Decision Tree Fit Statistics

Variable Importance

Variable Name	Label	Number of Splitting Rules	Importance	Validation Importance	Ratio of Validation to Training Importance
PMT		1	1	0.688654	0.688654
MSAT		1	0.506153	1	1.975688
RANK		1	0.424597	0.142487	0.335582
VSAT	VSAT	0	0	0	.
GENDER		0	0	0	.

Table 2. Variable Importance

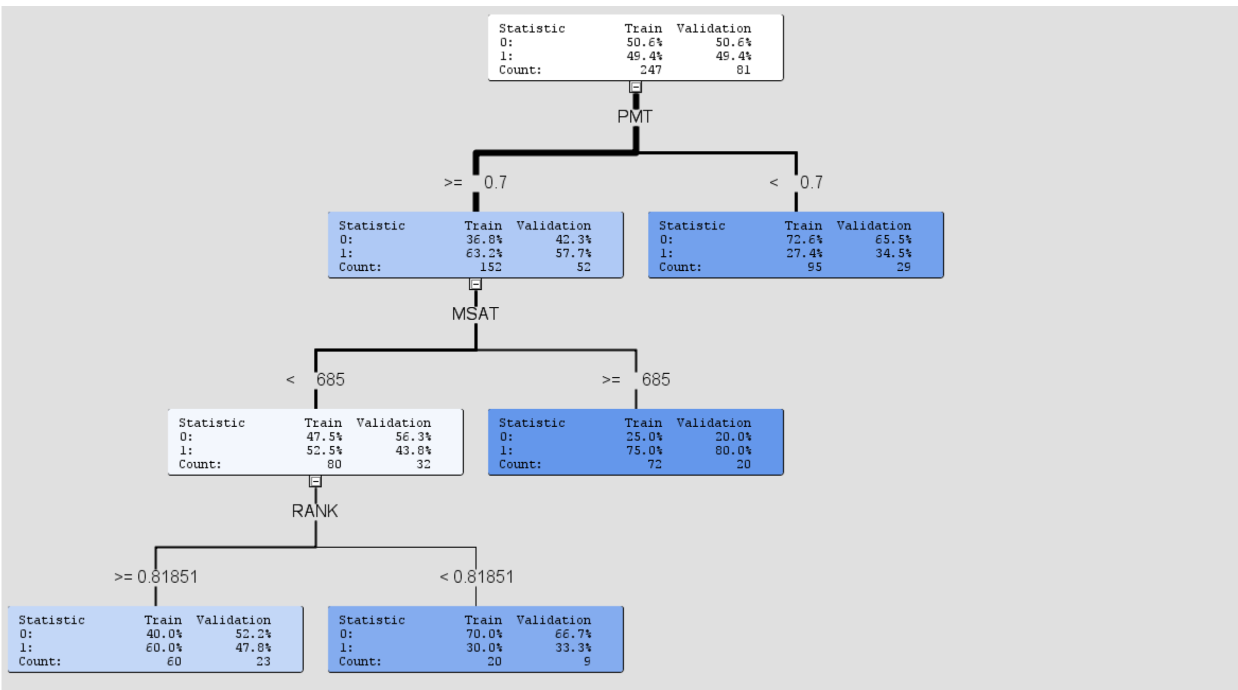


Figure 2: Decision Tree

It can be noted that the tree is explanatory and easy to interpret. We can see that our placement exam is most important in determining success of our actuarial students, followed by their Math SAT score and then finally by their rank in class in their senior year of high school.

When running the node with gradient boosting, we obtained the following results. We include tables of variable importance and fit statistics.

Variable Name	Label	Number of Splitting Rules	Importance	Validation Importance	Ratio of Validation to Training Importance	Interaction Importance
MSAT		2	1	0.79292	0.79292	0.097056
PMT		82	0.664983	1	1.503799	0.042872
VSAT	VSAT	1	0.41494	0.289852	0.698539	0.022006
RANK		7	0	0	.	0.016658
GENDER		0	0	0	.	.

Table 3. Variable Importance with Gradient Boosting

Interestingly, importance now has the ordering of MSAT, PMT, and VSAT and does not give much emphasis to RANK or GENDER. The tree by itself had the ordering of PMT, MSAT, and RANK and did not give much emphasis to VSAT or GENDER.

Target	Fit Statistics	Statistics Label	Train	Validation	Test
MAJOR	_NOBS_	Sum of Fre...	247	81	.
MAJOR	_SUMW_	Sum of Cas...	494	162	.
MAJOR	_MISC_	Misclassific...	0.336032	0.320988	.
MAJOR	_MAX_	Maximum A...	0.59521	0.625447	.
MAJOR	_SSE_	Sum of Squ...	118.3984	39.38001	.
MAJOR	_ASE_	Average Sq...	0.239673	0.243086	.
MAJOR	_RASE_	Root Avera...	0.489564	0.493038	.
MAJOR	_DIV_	Divisor for A...	494	162	.
MAJOR	_DFT_	Total Degre...	247	.	.

Table 4. Fit Statistics of Gradient Boosting

Finally, in comparing the results of all methods used in this study as well as the previous one, we note that the logistic regression had the lowest misclassification error, followed by gradient boosting. The gradient boosting method was better than the decision tree by itself. It should also be noted that the logistic regression would not have done as well except for the imputation that occurred in a prior operation. Otherwise, there would have been substantial loss of cases.

Fit Statistics

Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Valid: Misclassification Rate ▲	Train: Sum of Frequencies	Train: Sum of Case Weights Times Freq	Train: Misclassification Rate	Train: Maximum Absolute Error	Train: Sum of Squared Errors	Train: Average Squared Error	Train: Root Average Squared Error	Train: Divisor for ASE	Train: Total Degrees of Freedom	Valid: Sum of Frequencies	Vali of C Weig Time
Y	Reg	Reg	Regression	MAJOR	0.283951	247	494	0.336032	0.8995	105.3901	0.21334	0.461888	494	247	81	
	Boost	Boost	Gradient Bo...	MAJOR	0.320988	247	494	0.336032	0.59521	118.3984	0.239673	0.489564	494	247	81	
	Ensmbl	Ensmbl	Ensemble	MAJOR	0.345679	247	.	0.294355	0.815904	94.60941	0.191517	0.437627	494	.	81	
	Neural2	Neural2	Neural Net...	MAJOR	0.345679	247	494	0.271255	0.958775	89.43514	0.181043	0.425491	494	247	81	
	Neural	Neural	Neural Net...	MAJOR	0.358025	247	494	0.327935	0.857516	100.1899	0.202814	0.450348	494	247	81	
	Tree	Tree	Decision Tr...	MAJOR	0.358025	247	494	0.299595	0.75	101.9684	0.206414	0.454328	494	247	81	

Table 5. Fit Statistics of Alternative Methods

CONCLUSIONS

Although logistic regression, neural networks and decision trees are popular methods for predicting a categorical variable, decision trees are probably a better choice because they generate transparent rules that are easily interpretable, especially by non-statisticians. Using gradient boosting can make decision trees more accurate in terms of reducing the misclassification rate.

However, there are some caveats. Note that the procedure for applying this method to classification problems requires that separate sequences of (boosted) trees be built for each category. Therefore, it is not wise to analyze categorical dependent variables with many classes as the computations performed may require an unreasonable amount of time. Trees themselves also have shortcomings. The description given by the relationship in the tree may not be the only accurate one. It might appear that certain inputs uniquely explain the variations in the target. However, a completely different set of inputs might give a different explanation that is just as accurate. Obviously, this was a very limited case and an extensive simulation would need to be conducted in order to make generalizations.

REFERENCES

[1] Applied Analytics Using SAS Enterprise Miner 5.3: Course Notes (2009).

[2] Friedman, Jerome H. (2001), “Greedy function approximation: A gradient boosting machine,” *The Annals of Statistics*, Vol.(29), 1189-1232. Available at http://stat.stanford.edu/~jhf/ftp/trebst.pdf.

[3] Friedman, Jerome H. (2002), “Stochastic gradient boosting,” *Computational Statistics & Data Analysis*, Vol.(38), 367-378. Available at http://stat.stanford.edu/~jhf/ftp/stobst.pdf.

[4] Schumacher, P., Olinsky, A., Quinn, J. (2010), “A Comparison of Logistic Regression, Neural Networks, and Classification Trees Predicting Success of Actuarial Students,” forthcoming in the *Journal of Education for Business*, Vol. (85), No. 5.

[5] SAS Enterprise Miner (2009), All output and data analysis for this paper were generated using SAS Enterprise Miner software, Version 6.1 of the SAS System for Windows. Copyright © 2009 SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

THE EFFECT OF LOGISTICS OUTSOURCING ON PLANT DELIVERY PERFORMANCE: AN EMPIRICAL STUDY OF THE U.S. AUTOMOTIVE SUPPLY CHAIN

Mary J. Meixell, Quinnipiac University, (203) 582-5206, mary.meixell@quinnipiac.edu
George Kenyon, Lamar University, (409) 880-8388, george.kenyon@lamar.edu

ABSTRACT

Logistics outsourcing has a significant effect on how manufacturing firms produce and deliver products to their customers. Indeed, many manufacturing firms do not own or manage the transportation and warehousing resources used for inbound and outbound shipments from their facilities. However, earlier research has demonstrated mixed results when outsourcing; some companies experience favorable performance outcomes while others do not. In this research, we investigate the effects of logistics outsourcing on delivery performance by analyzing empirical data from the automotive industry using data from a survey of manufacturing plant managers. Our analysis indicates that logistics outsourcing does not directly effect on-time delivery rate, but several moderating factors emerged that suggest interesting strategies for outsourcing the logistics functions.

Keywords: outsourcing, transportation, warehousing, third party logistics, automotive

INTRODUCTION

Firms seek to improve performance in cost as well as quality and delivery through the outsourcing of components, activities, and processes. Yet, earlier research has determined that plants have not experienced consistent improvement in performance when outsourcing some portion of their operation. Studies show that while many plants improve performance [1, 2], others have mixed results in terms of cost efficiency, productivity and profitability [3-5]; experienced a loss of innovation due to the decoupling of product research and development from production [6, 7]; loss of proprietary knowledge [8]; and, degradation in purchasing power due to volume loss [9]. Indeed, managers perceive this uncertainty in expected outcome that is associated with outsourcing as a risk in their supply chains [10].

In this study, we investigate the influence of logistics outsourcing on delivery performance, measured as on-time delivery rate. We focus on two activities that are often outsourced in the automotive supply chain – transportation and warehousing. The performance metric selected for analysis is the on-time delivery rate, due to its relevance in the automotive industry. Schedules across the tiers in automotive supply chains have long been tightly coordinated, especially at the tiers closest to the end customer, because inventory is not a viable method for a finished product with so many permutations of the end product. On-time delivery is also interesting in this context because suppliers come from a wide variety of industries. The top tiers in the supply chain include suppliers of stamped sheet metal, cast engine blocks, injected molded plastic, case alloy wheels, wiring harnesses, to electronic control modules. This presents variety in manufacturing environment, while imposing an enterprise relationship that is remarkably consistent from firm to firm.

We examine performance in logistics outsourcing in concert with practices that are likely to influence the outcome – strategy, structure and enterprise integration. Structure, and its alignment with product, has long been a subject of interest in operations management, but has not received as much attention in logistics and supply chain management [11]. Early research in operations strategy has investigated the relationship between strategy, structure and performance. Certainly the product-process matrix forms an important element of structure in the manufacturing industries, as proposed by Hayes and Wheelwright [12] and subsequently adopted into increasingly sophisticated structure models [13, 14]. Stock, Greis and Kasarda incorporated the logistics activities into this framework, and proposed that practices such as logistics integration can interact with strategy and structure choices to affect performance [15]. Here, we include product-process variable, structure, and enterprise integration into the model.

We investigate the relationship between logistics outsourcing and delivery performance. Specifically, we examine the following research questions:

1. How does transportation and warehousing outsourcing influence on-time delivery?
2. Are some industries more successful at this than others? In particular, have these types of outsourcing been successful in the automotive supply chain?
3. What contextual factors influence these results?
4. What mitigating factors influence these results?

EMPIRICAL ANALYSIS

To accomplish these aims, we use data from an annual survey of plant-level data on manufacturing metrics, management practices, and financial results. This data is collected by the Manufacturing Performance Institute (MPI), under the auspices of the Industry Week (IW) trade publication. According to the 2002 U.S. Census report there are 350,829 U.S. based manufacturing plants. In this data, we select and analyze the data reported by plants that participate in the automotive industry supply chain. For the independent variable of On-time Delivery Rate (OTD), the SAS Power procedure for determining sample size versus confidence interval half-width was used to determine the minimum sample size needed. Given a normal distribution with a two-sided test, an alpha (significance level) of 0.01, a desired probability of achieving the target precision of 0.95, a target half width of 1.00, and an estimated standard deviation of 9.4887, the minimum required sample size is 281. Thus, the levels of precision for this survey are reasonable for making inferences on the behavior of the studied population. The responses for IW/MPI survey data from 2004 through 2007, numbered 681, 668, 798, and 431 for each year respectively; yielding a total sample of 3441 responses. Of these respondents, 336 identified themselves as participants in the automotive supply chain (value chain), and 298 provided OTD rate data.

This survey contains over a hundred variables that pertain to how the responding plant structures its operation and utilizes its capacity, equipment, information technology, human resources, and supply chain. The primary and transformed variables utilized in this study are presented in Table 1, along with their summary statistics.

Table 1: Dependent Variables Utilized and Summary Statistics.

Dependent Variable	n	Mean	Std Dev	Min	Max	Skewness	Kurtosis
ON_TIME_DELIVERY_RATE	298	94.6455705	9.48870346	3	100	-4.9114648	35.3608727
LOG_ON_TIME_DELIVERY_RATE	298	3.26445076	1.12286497	-3.19867	4.61512	-1.1609184	3.14625584

The primary dependent variable is defined as follows: OTD = the plant’s on-time delivery rate percentage of all deliveries (Range: 0 to 100). The logit transformation is used to dramatically lessen the degree of non-normality seen in the OTD data [16].

The primary independent variables of interest in this study are the logistics outsourcing variables – i.e. whether or not the plant retained the operation and/or management of its transportation activities (OutTrans), and warehousing and distribution activities (OutWrhs), or if these functional activities were outsourced.

We also include three context variables in this study. The product volume and mix (Volmix) variable is used here as a proxy measure for the amount of logistics work at each respondent plant. Volmix indirectly provides definition on the processing strategy used by the plant; capturing to some degree the manufacturing cycle time difference between the various production strategies. The competitive environment component of the model is defined using the variable "Year" to capture environmental effects related to contextual impacts associated with each of the years that data was collected. Note that the use of a single supply chain also contributes to uniformity of the operational setting in this study.

The structure variable captures context relating to the supply chain and continuous improvement programs for the plant and market conditions. This variable is defined using the following IW/MPI variables: Who the final products are shipped to, Agile, Total Quality Management (TQM) systems, Degree of Supplier Integration, Degree of Customer Integration, Supplier Relations, Customer Relations, Percentage of Overseas Sales, and the Percentage of Imports. Cluster analysis was used to group the respondents into organizational groupings defined as; Narrow Hierarchy Structure, Narrow Market Structure, Narrow Network Structure, Wide Hierarchy Structure, Wide Market Structure, and Wide Network Structure.

Finally, moderating variables were included in this study to establish the conditions that are conjectured to influence success when outsourcing logistics. A key moderating variable is the strategy, which is defined using the strategic priorities variables captured in the survey. The MPI/IW survey asked each respondent to identify the "three objectives that best describe the focus of your market strategy," from a list of the following: low cost, high quality, fast delivery, innovation, product variety, customization, service and support, and total value. Cluster analysis was again used to classify respondents into four strategic groupings, consistent with earlier classification schemes in the literature [17]: Prospectors, Analyzers, Low Cost Defenders, and Differentiated Defenders. The last two moderating variables relate to whether or not information is readily shared and processes are integrated across enterprise lines, using the survey items that represented the degree of supplier integration (Supp-Int) and degree of customer integration (Cust-Int).

RESULTS

The results of this analysis, which were developed using the SAS ANOVA procedure, are summarized Table 2. First, we see that the OTD model was significant (Pr>0.0001) with an R-Squared value of 0.7142. The main effects of outsourcing transportation and warehousing on OTD were not, however, found to be significant (Pr>0.2267) and (Pr>0.1300), respectively. Customer integration has a strong influence on OTD (Pr>0.0001) independent of logistics outsourcing: an extensive level of integration versus some degree of integration has a difference of 0.5454; extensive versus none has a difference of 0.6523. Also, some of the strategies lend themselves to improved OTD: Low cost defender strategy, differentiated defender strategy and prospector strategy all have significantly better OTD results than the analyzer strategy. There are several significant interactive effects that are pertinent to firms that outsource logistics and participate in automotive supply chains that bear consideration, each of which is discussed in this section.

Table 2 shows that several of the main effect variables are significant. Both the structural alignment of the plant and its strategic orientation towards the marketplace significantly affect OTD, as well as the interactions between these two variables. Furthermore, the plant's production strategy and the degree to which it has integrated processes with customers significantly affect OTD.

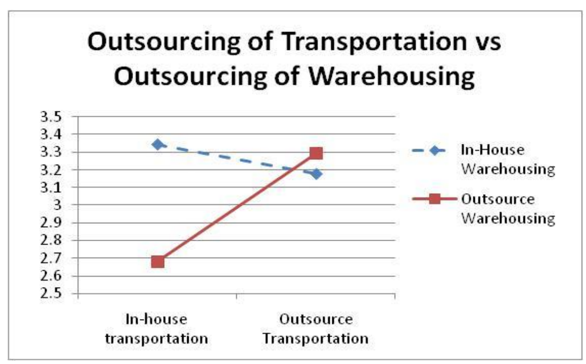
TABLE 2: ANOVA Results for Ontime Delivery Rate Model

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	66	225.6223827	3.4185209	5.72	<.0001
Error	200	119.5113718	0.5975569		
Corrected Total	266	345.1337544			

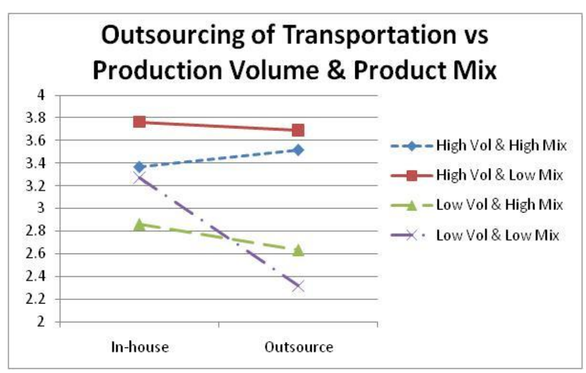
R-Square	Coeff Var	Root MSE	Log Ontime Delv Rate Mean
0.653724	23.83039	0.773018	3.243833 (97.2%)

Source	DF	Anova SS	Mean Square	F Value	Pr > F
year	3	5.48867221	1.82955740	3.06	0.0293
Cust_Int	2	16.29598182	8.14799091	13.64	<.0001
OutTrans	1	0.78442568	0.78442568	1.31	0.2533
OutWrhs	1	1.23344353	1.23344353	2.06	0.1524
Strategy	3	15.40771517	5.13590506	8.59	<.0001
Structure	5	7.35160939	1.47032188	2.46	0.0344
Supp_Int	2	1.90494521	0.95247260	1.59	0.2057
VolMix	3	45.37190625	15.12396875	25.31	<.0001
Strategy*Cust_Int	5	11.55398296	2.31079659	3.87	0.0023
OutTrans*Cust_Int	2	6.19389990	3.09694995	5.18	0.0064
OutTrans*Supp_Int	2	5.54649236	2.77324618	4.64	0.0107
OutTrans*Strategy	3	6.75337822	2.25112607	3.77	0.0116
Supp_Int*Cust_Int	3	6.69873002	2.23291001	3.74	0.0121
Strategy*Structure	9	11.71139242	1.30126582	2.18	0.0250
OutTrans*OutWrhs	1	2.99814032	2.99814032	5.02	0.0262
Strategy*Supp_Int	5	7.26019858	1.45203972	2.43	0.0364
OutTrans*VolMix	3	5.16800576	1.72266859	2.88	0.0369
OutWrhs*Supp_Int	2	3.50207773	1.75103887	2.93	0.0557
OutWrhs*Cust_Int	2	3.27606438	1.63803219	2.74	0.0669
OutWrhs*Strategy	2	28.28705225	14.14352613	23.67	<.0001
OutWrhs*VolMix	2	13.93398381	6.96699190	11.66	<.0001
Strategy*VolMix	5	18.90028468	3.78005694	6.33	<.0001

Of equal, if not greater, interests are the interactive effects between the independent variables. The interaction between outsourcing of warehousing and outsourcing of transportation shows that there is little influence on OTD when warehousing is outsourced, as the OTD is in the range of 96% for both cases. However, the outsourcing of transportation and warehousing together result an OTD increase from 93.7% to 96.4%. Thus, companies that outsource both transportation and warehousing benefit.

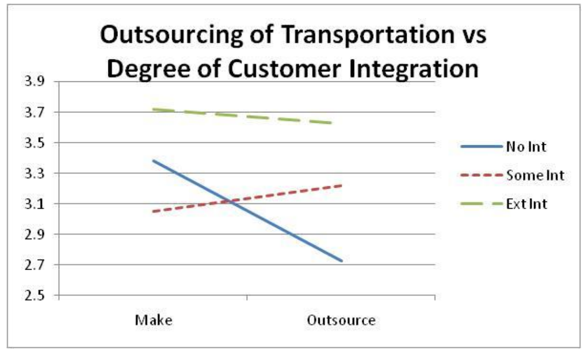


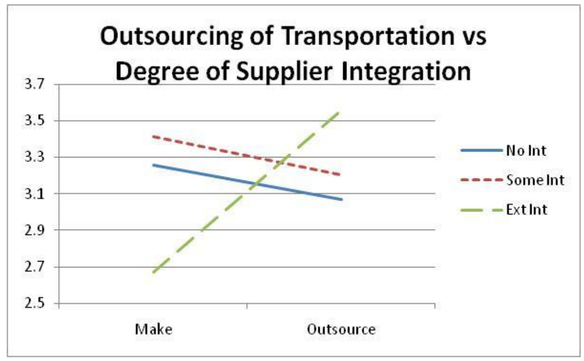
The interaction between transportation outsourcing and volume-mix is also significant, and suggests an important contextual factor. We see here that transportation outsourcing has a comparable effect on delivery performance for high volume (regardless of mix) and for high product mix (regardless of volume) environments. The high volume high mix environment has the highest OTD outcome regardless of whether

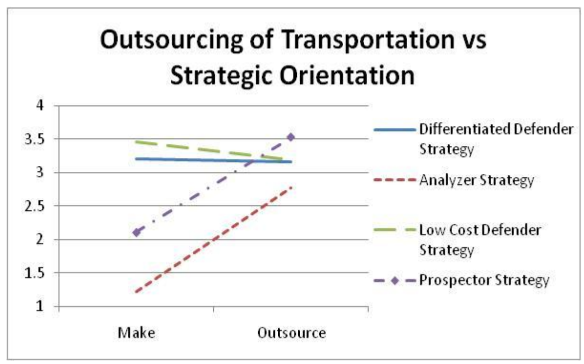


transportation is outsourced or managed in house. However, for the low volume, low mix environment, there is a dramatic fall in delivery performance (from 96.4% to 90.9%) when transportation is outsourced. Accordingly, when both production volume and product mix are low, OTD tends to drop for firms tend who outsource transportation. The interaction between outsourcing of warehousing and volume-mix is also significant, showing a significant degradation of outsourcing (approximately 7%) for both low volume –high mix and high volume – low mix.

The interactions between transportation outsourcing and integration are also significant and suggest mitigating strategies. Note that delivery performance falls from 97.3% to 93.7% when manufacturers outsource transportation with no customer integration. The effect is different, however, for transportation outsourcing with supplier integration. Here, the firms in our sample saw a notable increase in OTD (from 93.7% to 97.1%) when done in conjunction with extensive supplier integration. As a result, companies that don't integrate their processes with their customers will have significantly lower OTD when outsourcing transportation. On the other hand, firms that have extensive supplier integration will tend to improve OTD when outsourcing transportation.







Finally, the analysis also suggests that the plant's strategic orientation is important to OTD when outsourcing transportation. Here we see little effect in OTD in firms with a differentiated defender or low cost defender strategy – both types of firms have an OTD around 96% regardless of whether they outsource transportation or keep it in-house. The outcome of outsourcing is notably different with firms that follow an analyzer or prospector strategy. Here, outsourcing transportation has a sizeable improvement in OTD, from 77.8% to 94.0% for analyzer firms, and from 89.1% to 97.1% for prospector firms. As a result, firms with analyzer or prospector strategies tend to improve their OTD greatly when outsourcing transportation.

CONCLUSIONS

In this research, we investigated the effects of logistics outsourcing on plant level delivery performance. We draw on plant-level performance data collected by the IW/MPI Census of Manufacturers survey, and build statistical models to aid in understanding the degree to which firms benefit from this type of outsourcing practice, as well as the conditions under which it is most likely to fail or succeed. The results suggest that neither transportation or warehousing outsourcing alone significantly influences on-time delivery performance, but that there are mitigating conditions as follows: 1) companies that outsource both transportation and warehousing benefit, 2) companies that don't integrate their processes with their customers will have significantly lower OTD when outsourcing transportation, 3) firms that have extensive supplier integration tend to have an improvement in OTD when outsourcing transportation, 4) when both production volume and product mix are low, firms tend to do worse in OTD when transportation is outsourced, and 5) firms with analyzer or prospector strategies tend to improve their OTD greatly when outsourcing transportation.

REFERENCES

1. Lei, D. and M. Hitt, *Strategic restructuring and outsourcing: The effect of mergers and acquisitions and LBO's on building firm skills and capabilities.* Journal of Management, 1995. **21**(5): p. 835-859.
2. Quinn, J.B., *Intelligent enterprise: A knowledge and service based paradigm for industry*. 1992, New York, NY: Free Press, Inc.
3. Ehie, I.C., *Determinants of success in manufacturing outsourcing decisions: A survey study.* Production and Inventory Management Journal, 2001. **42**(1): p. 31-39.
4. Jiang, B., G.V. Frazier, and E.L. Prater, *Outsourcing effects on firms' operational performance: An empirical study.* International Journal of Operations and Production Management, 2006. **26**(12): p. 1280-1300.
5. Kotabe, M. and J.Y. Murray, *Global sourcing strategy and sustainable competitive advantage.* Industrial Marketing Management, 2004. **33**(1): p. 7-14.
6. Fifarek, B.J., F.M. Veloso, and C.I. Davidson, *Offshoring technology innovation: A case study of rare-earth technology.* Journal of Operations Management, 2008. **26**(2): p. 222.
7. Bengtsson, L. and C. Berggren, *The integrator's new advantage - The reassessment of outsourcing and production competence in a global telecom firm.* European Management Journal, 2008. **16**: p. 635-643.
8. Wu, F., H.Z. Li, L.K. Chu, and D. Sculli, *An outsourcing decision model for sustaining long-term performance.* International Journal of Production Research, 2005. **43**(12): p. 2513.
9. Ellram, L. and C. Billington, *Purchasing leverage considerations in the outsourcing decision.* European Journal of Purchasing & Supply Management, 2001. **7**(1): p. 15-27.
10. Tang, C.S., *Perspectives in supply chain risk management.* International Journal of Production Economics, 2006. **103**(2): p. 451.
11. Grover, V. and M.K. Malhotra, *Transaction cost framework in operations and supply chain management research: Theory and measurement.* Journal of Operations Management, 2003. **21**(4): p. 457-473.
12. Hayes, R.H. and S.C. Wheelwright, *Link manufacturing process and product life cycles.* Harvard Business Review, 1979. **57**(1): p. 133-140.
13. Devaraj, S., D.G. Hollingworth, and R.G. Schroeder, *Generic manufacturing strategies: An empirical test of two configurational typologies.* Journal of Operations Management 2001. **19**: p. 427-452.
14. Kotha, S. and D. Orne, *Generic manufacturing strategies: A conceptual synthesis.* Strategic Management Journal, 1989. **10**: p. 211-231.
15. Stock, G.N., N.P. Greis, and J.D. Kasarda, *Enterprise logistics and supply chain structure: the role of fit.* International Journal of Physical Distribution and Logistics Management, 1999. **29**(4): p. 37-52.
16. Kutner, M.H., C.J. Nachtsheim, and J. Neter, *Applied Linear Regression Models*. 4th ed. 2004, Boston, MA: McGraw Hill.
17. Olson, E.M., S.F. Slater, and G.T.M. Hult, *The importance of structure and process to strategy implementation.* Business Horizons, 2005. **48**: p. 47-54.

Port Measurement Issues

In addressing how global ports will compete for the predicted re-shifting of container volume, traditional measures focus on attracting new container volume and retaining existing volume. Totally separate measures address how well ports attract new investment to improve port infrastructure, superstructure, and facilities. The problem with these measures is that they fail to capture the interaction effects between capturing new business and new investment. This suggests that ports may indulge in activities to improve one factor while harming the other. The purpose of this study is to use a new scale that measures these interaction affects and becomes the first to predict the impacts of the decisions of port administrators on customers and investors.

The general Research Question 1 is the following:

> *What are the factors on which ports compete? How can port administrators predict the impact of their decisions regarding port infrastructure as to become and remain competitive?*

The Changing Competitive Environment

From the time of the pioneer sailing on April 26, 1956 of Ideal-X, the first containership, from Newark, New Jersey, international trade and transport have been transformed through containerization. The result has been an alteration of the geography of production and distribution[1], with Asia becoming the global center for all types of manufacturing and service creation. Starting with Singapore, Hong Kong, and Busan, many Asian ports have taken full advantage of containerization and associated inter-modal transport methods to create some of the world's most efficient and competitive ports (Yap and Lam 2006).

In recent years, Asian merchandise trade has grown at a rate higher than the rate of growth of global trade in North America and Europe (see Table 1). Within Asia, of course, the main driver of merchandise trade is China where imports as well as export growth rates in recent years have hovered around 25 percent per year (see Table 1). While merchandise trade growth rates are high for Europe, trade among the European Union member states accounts for most of this growth. Asian trade, on the other hand, is directed significantly outside the region although intra-Asia trade is also growing. Table 2 further dramatizes the very high levels of North America and Europe-linked merchandise trade from China and Korea. In 2005, Asia's North America and Europe-linked trade was about $1.1 trillion dollars while North America's and Europe's combined Asia-linked trade was about $600 billion dollars.

Table 1: Share in Merchandise Trade by Global Regions or Countries

Region or Country	Exports	Imports

[1] The preceding sentences are based on 2007 Conference CFP "Globalization and Freight Transportation in a Containerized World", available at: http://www.gfptt.org/Entities/EventProfile.aspx?list=all&id=5621f7a8-eae8-4847-8e34-a1499c156391, Accessed on March 1, 2007.

	1990	2000	2005	Percent Change 2000-5	1990	2000	2005	Percent Change 2000-5
North America	16.6	19.5	14.5	4%	19.6	25.8	21.7	6%
Europe	49.6	42.0	43.0	11%	50.1	42.4	43.2	10%
Asia (Korea, China)	21.8	26.4	27.4	11%	20.3	22.9	24.7	12%
- China	1.8	4.0	7.5	25%	1.5	3.4	6.3	24%

Source: Based on World Trade Organization statistics, available at: http://www.wto.org/english/res_e/statis_e/its2006_e/its06_bysubject_e.htm , accessed on: September 10, 2007.

Table 2: Merchandise Trade among North America, Europe, Asia Regions

		Merchandise Trade with ($billions, 2005):		
		North America	***Europe***	***Asia***
Merchandise Trade from ($billions, 2005)	***North America***	824	238	270
	Europe	398	3201	332
	Asia	608	498	1424

Source: Based on WTO statistics, available at: http://www.wto.org/english/res_e/statis_e/its2006_e/its06_bysubject_e.htm , accessed on: September 10, 2007.

Asia's merchandise trade with North America and Europe is increasing faster than the planned capacity expansion for ports located in those regions. Exacerbating the imbalance problem is the fact that for the first time in many centuries, new shipping channels are being created and the usefulness of existing channels are being expanded beyond normal operating seasons. For example, several shipping companies are considering using the Northwest Passage across the top of the world as an alternative to the traditional Pacific routes. Reportedly, this will cut shipping time from Asia to the US east coast by 10%. Global warming effects have allowed the Northeast channels and seaports to operate later in winter season. As new lanes open up the probability of supply chain interruptions increase. This is somewhat counter-intuitive since more channels and extended usefulness should logically translate into improved balance. Instead these changes have led to greater imbalance due to the future ability of the Panama Canal to handle container cargo that was once transshipped through Busan and Los Angeles/Long Beach, reducing the demand on those infrastructures and the associated need for existing capacity and return on premium investments. This activity then increases the business at US east coast ports such as Houston, New Orleans, and Savannah; as well as European ports such as Hull in Great Britain and LeHarve in France. Longer shipping seasons through northern-most channels also reduce the need for developed premium hinterland and seaport capacities, which investors relied on to earn returns. To mitigate these emerging imbalances and potential interruptions, several initiatives have been undertaken, or are being contemplated (Song & Yeo, 2004; Kerr 2006; Anderson et al, 2007, Grigalunas, et al, 2009, et al, 2007; Jacobs & Hall, 2007: G). They include:

- Extreme pressure on the Long Beach port in California is spawning initiatives to enhance East Coast and Gulf ports in USA, and to increase volumes to Canadian ports, with surface transport of some of the goods to US.

- China’s development push in its own western region has shifted container volume to inland Yangtze ports - with Shanghai as a hub. This is necessary since 90% of China’s international trade is handled through marine transport.

- The intense transshipment competition between major Asian ports such as Busan, Singapore, and Shanghai has led to major investments in port capacities.

- Middle East ports such as Dubai Jebel Ali, Jeddah, are making continuing efforts and Aqaba to upgrade and to become transshipment centers not just for Middle East/Africa but also for South Asia – capturing some of the expected growth through Busan and China.

- China has launched aggressive plans to upgrade the infrastructure of its remote western regions by building a major highway link from its western provinces to the Pakistan port of Gwadar.

- Busan has consistently increased its berth capacity among its five member ports to discourage competition and maintain competitive service levels. This includes the development of the Busan New Port into a super-hub as well as the planned renovation of the existing Busan port. As part of this initiative, aggressive plans to increase hinterland access to the ports through the Siberian Railway are in the concept stages.

- The European Union, with the support of agencies such as Geneva-based International Road Transport Union (IRU), is making concerted efforts to improve transportation infrastructure by developing regular road-rail, inter-modal links, via northern and southern routes, to western China.

- To mitigate the increased costs and decreased service levels at some US ports, the Panama Canal widening project has prompted private firms to invest in increasing capacities on the US Gulf and East Coast ports such as Houston, Savannah and Charleston. In addition, CMA-CGM has chosen the Port of Boston as a transshipment port in the Northeast US. This compliments its Southeast transshipment port in Jamaica.

These initiatives are designed to mitigate the anticipated capacity imbalances that could potentially interrupt important supply chains, harm trade, and weaken national economies. This leads to the second research question:

Research Question 2

Will the existing actions by ports improve their competitive positions?

2. Literature Review

In addressing Research Question 1, the literature describes ten factors measuring two dimensions of port competitiveness. Included in these ten factors are variables measuring a port's ability to attract new investment. New investments in port infrastructure can lead to reduced congestion and provides resources to manage growth.

The literature on how ports compete has primarily addressed port competition from the viewpoint of the customer (customer-facing) within a country or region (e.g., Song & Yeo, 2004). Other studies either do not specify the context for their findings (i.e., whether the findings apply within a specific region or on a global scale) or they address only a few of the competitive factors predicting volume (e.g., Grigalunas et al, 2009; Brooks, 2000; Foster, 1979; Malchow and Kanafani, 2001; Murphy et al, 1988, 1992; and Slack, 1985). Collectively, these studies identified five major factors affecting port competition from the viewpoint of a customer. These include *port location, cargo volume, service level, port facility*, and *port expenses* (port cost or price). In the Song & Yeo (2004) model, the authors expressly exclude port expenses because the accounting practices in the Chinese ports they studied were heterogeneous – making meaningful comparisons difficult. In this study, the variable 'port costs' were included as a subjective factor so that Port Competition can be measured in a more robust manner. In doing so, it is expressly assumed that the port experts, managers, and directors have knowledge of these factors in the absence of objective data. Based on these studies, we adapted general definitions for each of the five competitive factors. Greater levels of each of the following factors are considered to make a port more competitive:

(1) Actual Cargo Volume (not capacity): Carriers and port users view major ports that handle large volumes of containers as preferable. Total volume combining import, export, and transshipment cargo appears to be more important than any single category of volume.

(2) Port Facility Capacity: Port facilities are defined as all tangible assets that are used to service water-borne cargo. Capacities of these assets are of particular importance to liners and carriers since ports operate during peak and off-peak periods. They include infrastructure, superstructure, and labor assets.

(3) Port Location: Location factors include geographical distance from production facilities, ease of port and berth entry, potential for expansion, and quality of inter-modal access.

(4) Service Level: The percentage of cargo that will be off-loaded/loaded within the port management's promised time period (variance of time promised), as well as the average off-load/load time. This includes operation during peak periods and in adverse weather conditions.

(5) Port Costs: The cost to the liner per TEU for load/off-load service as well as applicable port duties. From the viewpoint of a liner/user, Cost is a surrogate for port efficiency – so efficiency will not be measured separately (Song and Yeo, 2004).

Perhaps the most important issue with Asia-linked trade today entails the second factor listed above – Port Facility Capacity – because the existing 10,000, 14,000, and planned 18,000 TEU post-Panamax vessels require 16-21 meter, deep draft, multi-berth container ports (World Bank, 2007, Module 2, Pg. 41, Box 12). Currently, neither the US nor Busan (Korea) or China ports can handle all of these vessels when fully loaded (Haralambides and Behrens 2000). While these ports can handle these vessels when partially loaded, because the draft requirement is not as deep, they are appropriate for transshipment purposes. From the viewpoint of a customer/user, this could be especially problematic for developing countries because their fast growth and steady shift to world-class manufacturing will rely on developing reliable and efficient supply chains. The efficiencies created by these larger vessels will necessarily involve accommodating them when fully loaded. Not doing so will increase the probability of supply chain interruptions because of inadequate scale and efficiency of ports. Other well-known Asian-port rivalries (e.g., Busan-Shanghai, Singapore-Dubai) are motivating these mega-ports to improve their competitiveness for transshipment business by planning to build the capacity to handle the larger vessels. In addition, they are planning greater hinterland capacity to airports, roadways, and rail to improve the efficiencies of imports and exports.

In addition to the factors considered important to a customer (customer-facing competitiveness), the literature identifies factors affecting port competitiveness for resources (e.g. Foreign Direct Investment (FDI) and/or governmental funding). We group these factors under the label of Investment Competitiveness. These competitive factors are heavily influenced by public policy and could be used to shape public policy. [Source: World Bank Report, 2007] These factors are:

(1) Legal Framework: This entails the degree of autonomy of port management, including its own judiciary, to work outside of political arenas (similar to US Administrative courts) and the specific agreement between central and local governments describing the powers of port officials.

(2) Institutional Structure: The management structure should be conducive to investment (Private Sector ports), with proper autonomy and have a cooperative relationship with labor. The labor force should be sufficient and well trained. Table 3 demonstrates the four major port management structures and spells out how they differ with respect to the control over port assets and activities. They are Public Service Ports, Tool Ports, Landlord Ports, and Private Sector Ports. Table 4 describes each management structure – including its strengths and weaknesses. At the extremes, Public Service ports are operated as not-for-profit entities whose primary goal is public service; while Private Sector ports operate solely in the interest of the investors – with the government abdicating its rights for any public good.

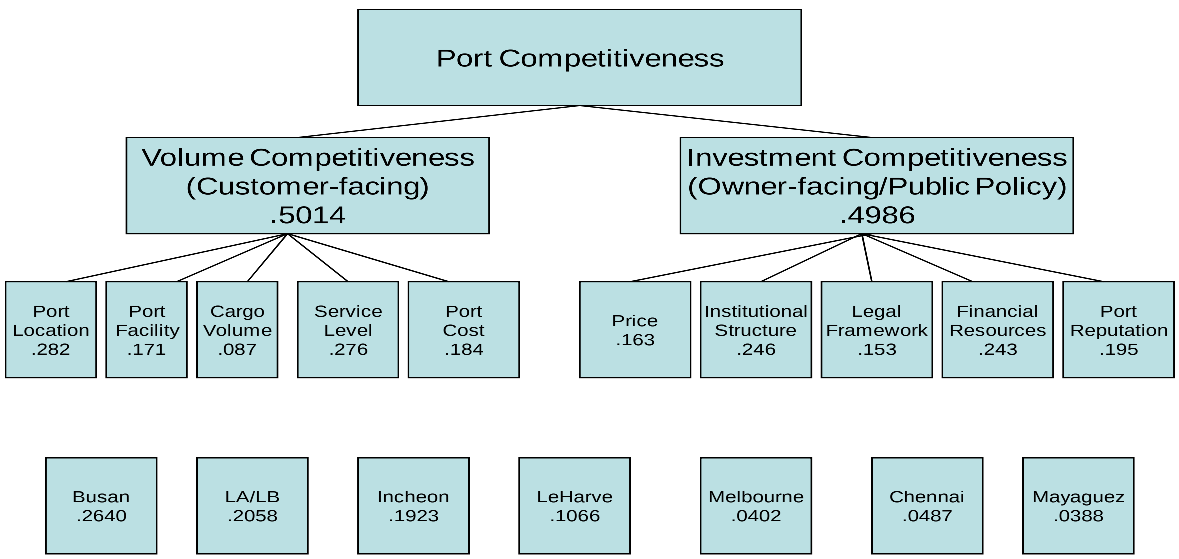
(3) Financial Resources: Autonomy to use port revenue for maintenance and expansion, maintain healthy cash flows, and the capacity to raise funds when needed.

(4) Port Reputation: The use of a mechanism, such as a port sector regulator, to ensure fair competition among the various entities that compete in ports. This involves preventing anti-competitive practices that often take place with port monopolies.

(5) Price: The price that ports charge for basic services including container handling, drayage services, premiums for peak periods, and storage fees.

The literature suggests that the public policy factors are key factors for a port's competitiveness for investment income. Investment sources include local private, FDI, local governmental, central governmental, and international agencies or intergovernmental. As described in Table 3, since all seaports have some accountability to government public policy can affect the competitiveness of all ports, regardless of purpose.

Results of Analytical Hierarchy Analysis (AHP) On The Sample Data



LATERAL TRANSHIPMENT WITH CUSTOMER SWITCHING

Wenjing Shen, Drexel University, (215) 895-0225, wenjings@drexel.edu
Xinxin Hu, Indiana University, (812) 855-3492, hux@indiana.edu
Yi Liao, Drexel University, (973) 342-5961, yl36@drexel.edu

ABSTRACT

We consider the inventory transshipment problem between two retailers that sell the same product to two independent markets. One assumption that distinguishes our work is that unsatisfied customers may also switch to the retailer with surplus inventory. We examine the impact of such customer switching behavior on firms' inventory decisions and profits. We identify situations when the firm with surplus inventory is willing to fully tranship, partially tranship, or not tranship. We show that there is a unique Nash Equilibrium in inventory decisions under certain conditions, and compare the case with central coordination with decentralized decision making.

INTRODUCTION AND LITERATURE REVIEW

When a customer walks into a retail store and finds the product she intended to purchase unavailable, both the retailer and the customer may take action. The retailer may request transhipment from a partner store where the product is available. If the retailer is not able to arrange the transhipment, the customer may turn to the partner store and buy the product by herself. The interaction of these two strategies and its impact on retailers' inventory decisions are interesting to study.

Inventory transhipment between retailers has been a common strategy when stockout occurs at one location and inventory surplus exists at the other location. Such strategy can be viewed as a way of virtually centralizing stock and taking advantage of inventory pooling. Therefore, when transhipment decisions are made under a central coordination scheme, it increases profit. However, if such a central coordination scheme does not exist, transhipment decisions are made locally for the interest of each individual retailer. As pointed out by Rudi et al. [4], such *interfirm* transhipment, rather than *intrafirm* transhipment, changes firms' inventory decisions since the risks imposed upon retailers are altered.

On the other hand, customer switching has been a well-observed phenomenon. If two retailers carry the same products, customers first go to their preferred retailer due to factors such as price, service, location, promotion, convenience, etc. Since the two retailers differentiate themselves in these aspects, they each have their own customer demand and do not compete directly. However, if customers cannot find the product in their preferred retailer, they may look for the product at another retailer. Not all customers do so since some of them may decide not to buy the product at all or buy a substitutable product. Such customer switching behavior is called “market search" in Anupindi and Bassok [2], where they assume that a fixed fraction of the customers will search for the goods at another retailer. However, since firms may not obtain the exact information on the proportion of customers who switch to another retailer, we consider in this paper that a random fraction of those will switch.

While making inventory purchasing decisions, it is important that firms take into consideration both transhipment agreement and customer switching behavior, since they interact with each other. In this paper, we examine two situations. In the centralized case, the two retailers are coordinated to maximize total profits. In the decentralized case, each retailer maximizes its own profit. In each case, we examine the optimal inventory purchasing decisions and transhipment decisions and aim at answering following questions:

- What is the optimal transhipment strategy for the firm with surplus stock? How much inventory to reserve and how much to tranship? What factors influence such decisions?
- In centralized case, what is the optimal inventory purchasing strategy? In decentralized case, is there a unique Nash Equilibrium for inventory purchasing quantities?
- How do the decisions in decentralized case compare to centralized case? Shall the firms order more or less inventory? Shall the firm with surplus reserve more or less inventory?
- Is there a coordination transhipment price such that centralized profit can be achieved under decentralized setting?
- What is the impact of customer switching behavior on inventory decisions, transhipment decisions, and profits? Should such switching behavior be encouraged?

Our project is closely related to the literature focusing on lateral transshipment in a decentralized system. The typical literature study the equilibrium behavior, and check whether it is possible to design a mechanism to coordinate those independent partners to achieve the highest possible supply chain benefit. Rudi et al. [4] and Hu et al. [3] show the uniqueness of Nash equilibrium in the order quantities for two retailers, and discuss the existence of the coordinating transshipment prices. Anupindi et al. [1] study the coorperative game in transshipment stage but non-cooperative game in order stage. A common assumption shared by these papers is that one partner's demand never switches to the other partner even if the local market stocks out but the other market still has inventory. Hence, if there is no transshipment between them, each partner's market is completely separated from the other's, and there is no competition on attracting demand. While in our paper, we assume that consumers in one market do “market search" by themselves and go to buy from the “remote" market when they find their local market is out of stock.

MODEL ASSUMPTIONS

We consider a one-period model where two retailers i, j= 1, 2 sell to their independent markets with random demand D_i, D_j, respectively. We label D_i, D_j as “local demand". Let $G_i(.)$ and $g_i(.)$ be the cdf and pdf for D_i. We assume that G_i and G_j are twice differentiable and strictly increasing on their supports. All the parameters for firm i and j can be asymmetric. All the assumptions below for firm i applies correspondingly to firm j.

Our model differs from Rudi et al. [4] by assuming that a random fraction of the unfulfilled demand after transhipment can switch to buy from another firm. The events take place as follows. Before observing demand, retailer i decides to purchase inventory Q_i. Simultaneously, firm j decides to purchase inventory Q_j. Then demand realizes and inventory is used to satisfy customer orders. Both demand and inventory decisions are observable to both firms. If one of the firms, say firm j, is short of inventory and the other firm, say firm i, has surplus, then firm j requests $q_{ij} = \min\{D_j - Q_j, Q_i - D_i\}$ from firm i. After q_{ij} units of inventory is transhipped to firm j, they are used to satisfy remaining customer demand. If there is still any remaining demand at firm j, then a random fraction A of the remaining demand will switch to firm i, which we label as “switching demand". We assume that A has cdf F(.), pdf f(.) and support [0, 1], and is independent of D_i and D_j. F is twice differentiable and strictly increasing. Firm i then uses its inventory available at this time to satisfy these switching customers.

Firm i pays unit cost c_i for each unit of inventory purchased before demand realizes, and receives revenue r_i for each unit sold to its own customers or to customers switching from firm j. Firm i receives p_{ij} from firm j for each unit of inventory transhipped from firm i to firm j, and pays τ_{ij} per unit as transportation cost. Any inventory left at the end of the period has salvage value s_i per unit. For each unit of unsatisfied local demand, firm i incurs penalty cost l_i. We assume that there is no penalty cost charged if firm i cannot satisfy switching demand. We believe that this assumption is reasonable since customers usually feel more disappointed with their retailer of first choice. We make the following assumptions: (1) $p_{ij} > \tau_{ij}+s_i$. (2) $p_{ij} < r_i+\tau_{ij}$. (3) $p_{ij} < r_j+l_j$. (4) $r_i> c_i$, $c_i < c_j + \tau_{ji}$, $s_i < s_j + \tau_{ji}$, and $r_i < r_j +\tau_{ji}$.

OPTIMAL DECISIONS UNDER CENTRAL CONTROL

In this section, we consider the situation when the decisions of the two firms are made under central coordination to maximize total profit. We first examine the transhipment decisions after demand at both firms are observed and satisfied using available inventory, then we examine the optimal inventory purchasing decisions. Assume now firm i has inventory surplus and firm j has inventory shortage. We make a decision on the amount of inventory q_{ij} to be transhipped from i to j in order to maximize total profit π^t. Define $x^+ = \min(x, 0)$.

LEMMA 1 When $D_i + D_j \le Q_i + Q_j$, if $r_j -\tau_{ij} + l_j -r_i EA -s_i (1-EA) \ge 0$, then $q_{ij}^t = D_j - Q_j$. Otherwise $q_{ij}^t = 0$.

Lemma 1 indicates that when surplus at firm i exceeds shortage at firm j, the optimal decision is either transhipping all needed inventory, or not transhipping anything at all. Because the latter scenario of never transhipping anything is not of interest for this paper, in the subsequent discussion we make the following assumption (A1): $r_j -\tau_{ij} + l_j -r_i EA -s_i (1-EA) \ge 0$.

When shortage at firm j exceeds surplus at firm i, depending on revenue and cost parameters, it might be more profitable to reserve some inventories at firm i anticipating that switching demand. Therefore an inventory rationing decision is needed. Let $z_i^t = (Q_i-D_i - q_{ij}^t)/(D_j - Q_j -q_{ij}^t)$, which represents the ratio of firm i's reserved inventory over the potential switching demand from firm j. Then the transhipment decision is equivalent to deciding z_i^t, i.e., to reserve how much percentage of the potential switching demand as inventory rather than shipping to firm j. For each z_i^t, $q_{ij}^t = [Q_i - D_i -z_i^t (D_j -Q_j)]/ (1-z_i^t)$.

LEMMA 2 If $D_i+D_j > Q_i + Q_j$, let z_i^{t*} be the solution of $\int_0^{z_i^{t*}} (1-a) f(a)\, da = (r_i+\tau_{ij} - l_j -r_j)/(r_i - s_i)$. Then if $(Q_i-D_i)/(D_j - Q_j) \le z_i^{t*}$, the optimal $q_{ij}^t=0$. Otherwise the optimal $z_i^t= z_i^{t*}$.

It is interesting to notice that z_i^{t*} is independent of D_i, D_j, Q_i, Q_j. Therefore, we should always try to keep the ratio of reserved inventory to potential demand as a constant. For example, if $z_i^{t*}= 0.2$, we should tranship some inventory such that the remaining inventory at firm i is 20% of the total potential switching demand. However, when inventory surplus at firm i is too small compared to shortage in firm j, it is possible that the remaining inventory is less than 20% of the total potential switching demand even without any transhipment. In this circumstance no transhipment should be arranged. Note that this transhipment decision is essentially a Newsvendor-type inventory decision. Therefore z_i^{t*} depends on the distribution of switching fraction A, as well as revenue and cost parameters. The following corollary summarizes these effects, which are intuitive.

COROLLARY 1 Given D_i, D_j, Q_i, Q_j, the optimal transhipment quantity q_{ij}^t is non-increasing in r_i, τ_{ij}, s_i and non-decreasing in l_j, r_j.

The optimal transshipment decisions can also be illustrated by Figure 1, with regions where only partial transhipment should be arranged, and regions where no transhipment should be arranged at all, even when firm j faces shortage and firm i has surplus inventory.

Next we consider the inventory purchasing decisions before demand realization. Under central coordination, we choose Q_i, Q_j to maximize total expected profit of the two firms. We can show that under certain conditions there exist a unique optimal pair of inventory decisions.

THEOREM 1 If $g_i(.)$ and $g_j(.)$ are non-increasing and all the parameters are symmetric, then $E\pi^t$ is jointly concave in (Q_i, Q_j). There exists a unique pair of (Q_i^t, Q_j^t) that maximizes expected total profit.

The condition of non-increasing density functions holds when demand follows uniform distribution or exponential distribution. Note that we do not impose any condition on the density function of the switching fraction f(.).

EQUILIBRIUM DECISIONS UNDER LOCAL CONTROL

In this section, we consider the case where a central coordination scheme to coordinate the inventory and transhipment decisions for the two firms does not exist. Therefore each firm acts on its own interest. We find that the policy structure bears great similarity to the centralized case.

LEMMA 3 When $D_i + D_j \leq Q_i + Q_j$, if $p_{ij} > \tau_{ij} + r_i\, EA + s_i\, (1\text{-}EA)$, then $q_{ij} = D_j\text{-}Q_j$. Otherwise $q_{ij} = 0$.

Similar to the centralized case, when total inventory exceeds total demand, either all shortage demand is fulfilled by transhipped inventory, or no transhipment takes place at all. However, the criterion is slightly different. For each unit of transhipped product, firm i receives p_{ij}, and pays τ_{ij} as shipping cost. Other than the shipping cost, firm i also incurs an opportunity cost since the expected revenue if otherwise reserved for switching demand is $r_i\, EA + s_i\, (1\text{-}EA)$. Therefore, firm i does not tranship any inventory is p_{ij} is less than the total expected marginal cost $\tau_{ij} + r_i\, EA + s_i\, (1\text{-}EA)$.

When total inventory is less than total demand, we define the same $z_i = (Q_i\text{-}D_i - q_{ij}) / (D_j - Q_j - q_{ij})$, which represents the ratio of firm i's reserved inventory over the potential switching demand from firm j.

LEMMA 4 If $D_i + D_j > Q_i + Q_j$, let z_i^* be the solution of $\int_0^{zi*} (1\text{-}a)\, f(a)\, da = (r_i + \tau_{ij} - p_{ij}) / (r_i - s_i)$. Then if $(Q_i - D_i) / (D_j - Q_j) \leq z_i^*$, the optimal $q_{ij} = 0$. Otherwise the optimal $z_i = z_i^*$.

Also similar to the centralized case, firm i should always keep the same ratio of reserved inventory over potential switching customers. The ratio z_i^* is different from z_i^{t*} under central coordination. Comparing these two ratios, we have the following observation:

LEMMA 5 For any given Q_i, Q_j, D_i, Dj, $z_i^* \geq z_i^{t*}$ and $q_{ij} \leq q_{ij}^t$.

Lemma 5 indicates that less inventory is transhipped without central coordination than with central coordination. This is due to the fact that shortage at firm j not only does not result in any penalty cost for firm i, but also generates more switching demand to firm i. Therefore, shortage at firm j is to a certain degree desirable to firm i. If firm i makes transhipment decisions to maximize its own profit, it would reserve more inventory and tranship less to firm j than if the two retailers are centrally coordinated.

A graphical representation of the transhipment decision is the same as Figure 1 except that all superscript t's are eliminated. Thus firm i may arrange fully transhipment, partial transhipment, or no transhipment depending on the amount of surplus compared to shortage.

Next we consider the inventory purchasing decisions if no coordination scheme exists. For any inventory purchasing decisions (Q_1, Q_2), the expected profit for firm i is given by:

$$E\, \pi_i(Q\_i, Q\_j) = -c_i\, Q_i + r_i\, E \min(D_i, Q_i) + (p_{ij}-\tau_{ij})\, Eq_{ij} + r_i\, E \min[(Q_i - D_i - q_{ij}), A\,(D_j - Q_j - q_{ij})]^+ + s_i\, E[(Q_i - D_i - q_{ij}) - A\,(D_j - Q_j - q_{ij})]^+.$$

We also show that under certain conditions a unique Nash Equilibrium exists.

THEOREM 2 If $g_i(.)$ and $g_j(.)$ are non-increasing functions, then there exists a unique Nash Equilibrium (Q_i^e, Q_j^e).

LEMMA 6 There does not exist a coordination transhipment price p_{ij} such that $\tau_{ij} + r_i\, EA + s_i\, (1\text{-}EA) \leq p_{ij} \leq r_j + l_j$ and $\pi_i\,(p_{ij}) + \pi_j\,(p_{ij}) = \pi^t\,(p_{ij})$.

Rudi et al. [4] show that there always exists a coordination price such that centralized profit can be achieved under decentralized control. Their results are later revised by Hu et al. [3] to show that such a coordination price only exists when the equilibrium inventory purchasing decisions for both firms are either more than or less than newsvendor quantity. However, when both transhipment and customer switching are considered, we show that such a coordination price does not exist in any situation. Rather than using a linear pricing mechanism, a more complicated mechanism is necessary for coordination and is interesting for future research.

SUMMARY AND DISCUSSION

In this paper, we consider a transhipment model with customer switching behavior. We consider both centralized and decentralized situations and study the inventory purchasing and transhipment decisions. If a firm with surplus inventory receives transhipment request from another firm with stockout, it may ship only part of the requested amount while reserving the rest in hand, since some of the unfulfilled demand at the stockout firm may switch to buy at the firm with surplus. Such strategic transhipment decisions also affect the inventory purchasing decisions before observing demand.

Our model leaves some interesting future research directions. First, how should the transhipment price be determined? Since the transhipment amount depends on price, a higher transhipment price results in less transhipped inventory, while a lower transhipment price results in more transhipped inventory. If one of the firms sets the price, there is an optimal price in between. It is unclear how endogenous transhipment price affects firms' inventory decisions. Secondly, is there a mechanism to coordinate the supply chain? And thirdly, it is interesting to observe that when partial transhipment takes place, higher shortage at firm j results in less transhipment from firm i. Therefore, firm j may have an incentive to underreport shortage. If shortage is unobservable, how would the shortage firm release information optimally? and how would the firm with surplus respond to that? Is there a mechanism that results in truth telling? All these questions are intriguing and we leave them for future work.

REFERENCES

[1] Anupindi, R., Y. Bassok, and E. Zemel. "A General Framework for the Study of Decentralized Distribution Systems," *Manufacturing & Service Operations Managemant*, 2001, Vol.(3), 349-368.

[2] Anupindi, R. and Y. Bassok. "Centralization of Stocks: Retailers vs. Manufacturer," *Management Science*, 1999, Vol.(45), 178-191.

[3] Hu, X., I. Duenyas, and R. Kapuscinski. "Existence of Coordinating Transshipment Prices in a Two-Location Inventory Model," *Management Science*, 2007, Vol.(53), 1289-1302.

[4] Rudi, N., S. Kapur, D.F. Pyke. "A Two-Location Inventory Model with Transshipment and Local Decision Making," *Management Science*, 2001, Vol.(47), 1668-1680.

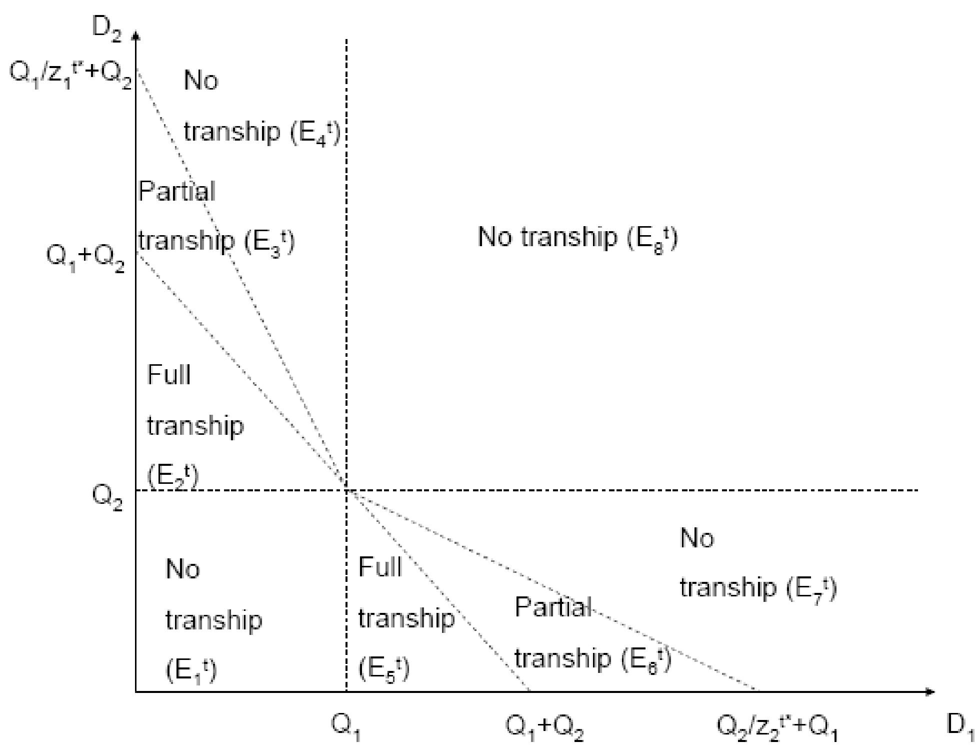


Figure 1: Transhipment Decisions Under Central Coordination

SALES INCENTIVES IN DISTRIBUTION CHANNELS: THE EFFECTS OF RETAILER'S PRICING SCHEME AND COMPETITION

Ozgun Caliskan Demirag
Sam and Irene Black School of Business, Penn State Erie, Erie, PA, 16563
ozc1@psu.edu

ABSTRACT

We analyze the effectiveness of sales incentives in multi-stage distribution channels where the retailer implements uniform pricing or price discrimination. Different forms of incentives by the manufacturers are analyzed such as threshold-controlled lump sum incentives and per-unit incentives, and coordination issues are discussed. In addition to the incentives given to the retailers, we analyze incentives directed to the customers, i.e., customer rebates. Incorporating the effect of competition, we also provide insights on the choice of promotion type by competing manufacturers.

Keywords: Retailer incentives, customer rebates, price discrimination, competition, automotive industry

INTRODUCTION

Many organizations are positioned in multi-stage distribution channels where they utilize retailers to sell their products to the end customers. This structure leads to indirect sales by manufacturers or wholesalers and allows different options for promoting their products. For example, a manufacturer may offer a sales incentive directly to the end customer in the form of a customer rebate, or may offer an incentive to its immediate customer (e.g., the retailer) in the form of a retailer incentive, both with the hopes of stimulating customer demand. In addition to the different choices of audience for directing the promotion, the manufacturer also has the flexibility to choose a particular structure for the promotion. For example, retailer incentives may sometimes be offered based on the size of the retailer's order quantity; in other cases there may be no strings attached to the promotion.

From the manufacturer's point of view, it is essential to know whether the promotions are effective in achieving the ultimate goal of improving sales and profits, and whether one particular form of promotion is superior to the others under certain conditions. Some important factors that influence the effectiveness of sales incentives are the retailer's pricing mechanisms and the existence of competition.

Motivated from the practices of automotive manufacturers in the United States domestic market, we investigate the effects of sales incentives on manufacturers' profits. In this industry, promotions are frequently offered, which may be given directly to customers or given to retailers (dealers) to increase demand. Retailer incentives can take different forms, e.g., lump-sum, per-unit, or quote-based incentives. One distinct characteristic of the retailers in this industry is that they are able to perform price discrimination (to some extent) due the nature of negotiation-based sales. Thus, it is interesting to investigate price-discriminating retailer as well as a uniform-price charging retailer. Competition is also an important component of the industry that may affect the companies' promotion decisions.

In our research, we investigate the performance of different sales incentives by incorporating issues such as price discrimination and competition. First, we analyze a single-manufacturer single-retailer channel and focus on retailer incentives that differ in terms of their dependence on the retailer's order quantity. In particular, we consider a (quantity-based) per-unit incentive and a (non-quantity based) lump-sum incentive scheme, where the retailer receives an incentive for each unit ordered in the former and a bonus payment independent of the order quantity in the latter. We analyze quantity-based incentive schemes where the retailer is qualified to receive a per-unit incentive (wholesale price discount) or a lump-sum payment only when the order quantity exceeds a threshold, the value of which is determined by the manufacturer. We consider two pricing strategies by the retailer. In one setting, the retailer charges a uniform price for each unit sold; in another setting, the retailer is equipped with the ability to exercise first-degree (perfect) price discrimination and may charge varying prices for each unit sold. Next, we analyze the case with two competing supply chains. We assume that the retailer is willing to sell a unit only if the selling price is above a certain value, which we refer as the retailer's reservation price. We study a single-period model where the manufacturer initiates the interactions by declaring the wholesale price and promotion amount and the retailer determines the order quantity. We use a downward-sloping linear demand function. We model the choice of promotion as an exogenous decision, hence we assume that the manufacturer commits to a certain type of promotion before engaging business relations with the retailer. We compare the manufacturer's profits under different incentives and also consider issues such as whether channel coordination is achieved with the use of incentives. In the case of competition, the manufacturers act simultaneously to determine their wholesale prices and promotion amounts, who are followed by retailers acting simultaneously to make their order quantity decisions. We model problems in a game theoretical framework and obtain insights using the Nash equilibrium decisions.

Marketing, economics and more recently the operations management literature have investigated the sales promotions and their roles on firms' profitability. Gerstner and Hess [6a] and Narasimhan [7], focus on the use of rebates in achieving price discrimination when there is no intermediary in the supply chain, and Gerstner and Hess [6b] introduce channel issues by into the analysis of the sales promotions. Bruce et al. [3a, 3b] analyze trade promotions (wholesale price discounts after a sales quantity target is met) and cash rebates by explicitly incorporating a durability measure for the manufacturer's products. There is a large amount of empirical research investigating how promotions work, focusing mostly on non-durables. See Blattberg et al. [2] for a review. In the operations area, Aydin and Porteus [1] and Chen et al. [5] are some examples of research in this direction. The first article compares rebates given to the customers and rebates given to the retailers modeling a stochastic demand. The latter focuses on the customer rebates and they show that they are effective in increasing the manufacturer's profits only when some customers fail to redeem their rebates.

Most relevant work is by Caliskan Demirag et al. [4], where the authors compare customer rebates and retailer incentives when the retailer can price discriminate. However, they do not consider threshold-based retailer incentives, which we include in our analysis. Different from their work, we also give insights on which of the two promotions, i.e., customer rebates and retailer incentives would work better for a manufacturer that is in competition with another manufacturer who also sells its products through a retailer.

MODELS

We present our model with threshold-controlled retailer incentives where the retailer is able to implement first-degree price discrimination. We model the setting as a two-stage distribution channel with a manufacturer and a retailer. The manufacturer determines the wholesale price *w*, and announces the type of retailer incentive along with the incentive value *K* (lump-sum amount) or *k* (per-unit amount), and threshold order quantity (*T*)). The retailer then decides the order quantity, which we denote with *Q*. Table 1 summarizes the notation.

Table 1: Notation

a :	Market potential
b :	Price sensitivity of customers
$P(Q)$:	Retail price when Q units are sold, i.e., $P(Q) = a - bQ$
Π^M :	Profit of the manufacturer
c :	Production cost of the manufacturer
w :	Wholesale price
$w + m$:	Reservation price of the retailer
K :	Lump sum incentive given to the retailer
k :	Per-unit incentive given to the retailer
T :	Threshold order quantity after which the retailer receives an incentive
r :	Retailer's per-unit price when uniform pricing is used

Due to the sequential decisions by the channel members, we solve for the players' equilibrium decisions using a backward induction procedure. We start with the retailer's problem when the manufacturer offers a threshold-controlled lump-sum incentive. The formulation is as follows. Given the manufacturer's decisions *w*, *K*, and *T*, the retailer solves the problems in (1) (where 0≤Q≤T) and (2) (where Q>T), and selects the order quantity that brings higher profit.

$$\begin{aligned} \max_{0\le Q\le T} \quad & \int_0^Q (a-bq)dq - wQ \\ \text{s.t.} \quad & Q \le \tfrac{a-w-m}{b} \end{aligned} \tag{1}$$

$$
\begin{aligned}
\max_{Q>T} \quad & \int_0^Q (a-bq)dq - wQ + K \\
\text{s.t.} \quad & \int_{\frac{a-w-m}{b}}^{Q} ((w+m)-(a-bq))dq \le K
\end{aligned} \tag{2}
$$

Then, we can show that the retailer is always better off by ordering at least T units, where his best response to the manufacturer's decisions is as follows:

$$
Q^*(w,K,T) = \max\left\{T, \min\left\{\tfrac{a-w}{b}, \tfrac{a-w-m+\sqrt{2Kb}}{b}\right\}\right\}. \tag{3}
$$

Anticipating the retailer's best response, the manufacturer solves the following problem:

$$
\begin{aligned}
\max \quad & (w-c)Q^*(w,K,T) - K \\
\text{s.t.} \quad & c \le w \le a-m \\
& \frac{(a-w-m-bT)^2}{2b} \le K \\
& T \ge \tfrac{a-w-m}{b}
\end{aligned} \tag{4}
$$

Optimizing the manufacturer's problem in (4), we obtain the manufacturer's best response. We summarize the equilibrium decisions in Proposition 1.

Proposition 1: When the manufacturer offers a lump-sum incentive to a perfectly price discriminating retailer for exceeding a threshold order quantity, the equilibrium is as follows: $w^* = a-m$, $K^* = \frac{(a-m-c)^2}{2b}$, $T^* = \frac{a-m-c}{b}$, $Q^* = \frac{a-m-c}{b}$, and $\Pi^M = \frac{(a-m-c)^2}{2b}$.

In the case with uniform pricing, the retailer chooses the per-unit retail price (*r*) to charge from every buyer. Continuing with the threshold-controlled lump-sum incentive, we formulate the retailer's problem as follows. The retailer solves the problems in (5) and (6), and selects the retail price that brings higher profit.

$$
\max_{r \ge \max\{w+m, a-bT\}} \quad (r-w)\left(\frac{a-r}{b}\right) \tag{5}
$$

$$
\begin{aligned}
\max \quad & (r-w)\left(\frac{a-r}{b}\right) + K \\
\text{s.t.} \quad & (w+m-r)^+\left(\frac{a-r}{b}\right) \le K \\
& r \le a-bT
\end{aligned} \tag{6}
$$

We can show that the retailer always finds it more profitable to use the incentive when provided. Let the retailer's best response be $r^*(w,K,T)$. Then, the manufacturer solves the following problem:

$$
\begin{array}{ll}
\max & (w-c)\left(\frac{a-r^*(w,K,T)}{b}\right)-K \\
\text{s.t.} & c \le w \le a-m \\
 & (w+m-(a-bT))T \le K
\end{array}
\tag{7}
$$

We summarize the equilibrium results in Proposition 2.

Proposition 2: When the manufacturer offers a lump-sum incentive to a uniform-price-charging retailer for exceeding a threshold order quantity, the equilibrium is as follows: $w^* = \frac{a+c-m}{2}$, $K^* = 0$, $T^* = 0$, $r^* = \frac{a+m+c}{2}$, and $\Pi^M = \frac{(a-m-c)^2}{4b}$ for $a \le 3m+c$, and $a-\sqrt{2m(a-m-c)} \le w^* \le a-2m$, $K^* = (w^*+m-(a-bT^*))T^*$, $T^* = \frac{a-m-c}{2b}$, $r^* = \frac{a+m+c}{2}$, and $\Pi^M = \frac{(a-m-c)^2}{4b}$ for $a \ge 3m+c$.

The procedure for finding the equilibrium decisions when the manufacturer offers a per-unit incentive after the retailer's order quantity achieves a threshold level follows similar steps to those shown above, hence they are omitted from this paper. We state our results for both forms of retailer incentives in Proposition 3. (For completeness we also list the previous findings by Caliskan Demirag et al. [4], where threshold-controlled incentives are not considered.)

Proposition 3: When the retailer exercises first-degree price discrimination,

(i) (Caliskan Demirag et al. [4]) both the lump-sum incentive and the per-unit incentive improves the manufacturer's profits,

(ii) the incentives further increase the manufacturer's profit when they are offered after the order quantity exceeds a threshold level. In this case, the equilibrium outcome brings the centralized channel profit, hence coordination is achieved.

When the retailer charges uniform price from all customers,

(i) (Caliskan Demirag et al. [4]) neither of the incentives improves the manufacturer's profit.

(ii) the incentives are able to increase the manufacturer's profit when they are offered after the order quantity exceeds a threshold level, but they are not able to achieve coordination in the channel.

In the case with competition, our focus is on comparing the retailer incentives and customer rebates. We model retailer incentive as lump-sum payments from the manufacturer to the retailer and rebates as per-unit payments from the manufacturer to the customer. The members of the supply chains interact as follows. The manufacturers move first and simultaneously make their wholesale price and promotion decisions. Next, the retailers observe the manufacturers' decisions and simultaneously make their order quantity decisions. Omitting the technical details in the analysis, we directly state our findings in Observation 1.

Observation 1: In the benchmark case with no promotions, we characterize the equilibria explicitly and identify the market conditions for a unique equilibrium to exist. We find that retailer incentives can be used by manufacturers to simultaneously improve each of their profits, but the retailers may have reduced profits as a result of the retailer incentive promotion. When manufacturers use customer rebates, we show that a manufacturer may be able to decrease the profit of her competitor while increasing her own profit, although she is also at risk for her competitor to use rebates similarly. We also observe from numerical examples that a manufacturer responds to the competitor's promotion with the identical promotion, except when the manufacturer has high production cost and sells to customers with high price sensitivity, in which case responding to the competitor's retailer incentive with a customer rebate can be more profitable.

CONCLUSION

Sales incentives are important promotional tools for organizations, which warrants the need for research investigating their effectiveness under different settings and scenarios. Focusing on promotions from the manufacturers to the retailers or customers and incorporating the considerations such as retailer's different pricing schemes as well as competition in multiple channels, we analyzed the effectiveness of the sales incentives on the manufacturers' profits and channel coordination. We further provided insights on the manufacturer's optimal choice of promotions in the presence of competition.

REFERENCES

[1] Aydin, G., E.L. Porteus. Manufacturer-to-retailer versus manufacturer-to-consumer rebates in a supply chain. Retail Supply Chain Management, Eds. N. Agrawal and S.Smith. Springer., 2008.

[2] Blattberg, R.C., R. Briesch, and E.J. Fox. How promotions work. Marketing Science, 14(3-2):G122-G132, 1995.

[3a] Bruce, N., P. Desai, R. Staelin. The better they are, the more they give: Trade promotions of consumer durables. Journal of Marketing Research, 42(1):54{66, 2005.

[3b] Bruce, N., P. Desai, R. Staelin. Enabling the willing: Consumer rebates for durable goods. Marketing Science, 25(4):350{366, 2006.

[4] Caliskan Demirag, O., O. Baysar, P. Keskinocak, J. Swann. The Effects of Customer Rebates and Retailer Incentives for a Manufacturer's Sales and Profits, Forthcoming at Naval Research Logistics, 2010

[5] Chen, X., Rhee B.-D. Li, C.-H., D. Simchi-Levi. The impact of manufacturer rebates on supply chain profits. Naval Research Logistics, 54:667-680, 2007.

[6a] Gerstner, E., D.M. Holthausen. Profitable pricing when market segments overlap. Marketing Science, 5(1):55-69, 1986.

[6b] Gerstner, E., J.D. Hess. Pull promotions and channel coordination. Marketing Science, 14(1):43-60, 1995.

[7] Narasimhan, C. A price discrimination theory of coupons. Marketing Science, 3(2):128-147,1984.

SUSTAINABILITY IN THE DUTCH ECONOMY FROM 1850 TO 2005

A.J.D. Lambert[1], TU Eindhoven, (+31)-040-2474634, A.J.D.Lambert@tue.nl
W.H.P.M. van Hooff, TU Eindhoven, W.H.P.M.v.Hooff@tue.nl
H.W. Lintsen, TU Eindhoven, H.W.Lintsen@tue.nl
J.L. Schippers, TU Eindhoven, J.L.Schippers@tue.nl
F.C.A. Veraart, TU Eindhoven, F.C.A.Veraart@tue.nl

ABSTRACT

The evolution of sustainability in the Dutch economy is studied over an extended period of time, i.e. 150 years. The study is based on Materials Flow Analysis which is complemented by analysis of qualitatively oriented documents. The research is mainly restricted to four years of reference, each marking the beginning of a new era in industrial development. The selected years are 1850, 1910, 1970, and 2005. These mark, respectively, the beginning of the Industrial Revolution based on coal, the introduction of electricity on a large scale, the emergence of environmental consciousness, and present time. Mass flows of international trade, domestic extraction, industrial production, consumption and waste management are combined, aimed at comparison of the sustainability in those years. Although the research is still in an initial phase, some preliminary conclusions can be drawn already.

INTRODUCTION

This research is part of the Dutch KSI-program (Knowledge network System Innovations) that aims the study of technological transitions in The Netherlands. It is focused on the generation of knowledge that might help to achieve transitions towards sustainability. Although qualitative analyses play an important role in these studies, these should be firmly underpinned with reliable quantitative data in order to generate valuable recommendations. To this order, the mass flows in the Dutch industry are studied into detail during four years of reference. These years have been selected via criteria of representativeness (not during war, recession or other non-typical events), and that they are at the verge of a new characteristic era. The years that have been selected are:

- 1850: Start of the industrial revolution based on coal
- 1910: Start of the 2nd industrial revolution based on electricity and mineral oil
- 1970: Emergence of environmental consciousness, start of information era, start of globalization
- 2005: Present

Materials Flow Analysis (MFA) has been adopted as an adequate instrument for this research [1]. This method considers the industry as a system and it is based on materials flows, which are expressed in mass units (kilotonnes). Only four distinct years are selected in order to create an as comprehensive as possible picture for every year. However, a drawback of this method is that even some long-term trends can become hidden as the sample years are far removed from each other. For instance, from the era of large-scale domestic coal mining only the start and the final phase are represented, by 1910 and 1970, respectively. The maximum reclamation of 5,883 kton in 1937 thus is not detected by this sampling and this example indicates that some concise study of intermediate years is desirable.

---

[1] Corresponding author

THE SYSTEM

The Netherlands is a small and specialized country with an open economy. The state, which initially included present-day Belgium, was established in 1813 after the defeat of the French emperor Napoleon Bonaparte. In the year 1830, however, the Belgian Uprising took place and Belgium declared itself independent of The Netherlands. The remaining part, which is present-day The Netherlands, was left with virtually no heavy industries. In Belgium, the industrial revolution based on coal and steam power started already about 1800 due to the work of William Cockerill, combined with the availability of coal and iron ore. On the Dutch territory, nearly all coal and iron had to be imported and, consequently, most industries were based on traditional methods. Therefore, only a few real industrial enterprises existed by 1850, such as the glass and pottery factory in Maastricht. Most important Dutch industries appeared not earlier than by the last quarter of the 19th century.

Although not an extremely small country, The Netherlands has never acquainted all possible branches of industry. Climate and geological circumstances caused the absence of some types of industry and the specialization on some other types, from which related branches emerged in turn. Examples of the latter were shipbuilding, pigment processing, and processing of luxury foods. Because of the modest size of the country, even when industrialization became stronger, particularly the heavy industries were represented with only one or a few plants. In this case, industrial history appeared helpful as we could include the year of erection and that of enclosure of some strategic individual companies. If, e.g., metallurgical industries are considered, The Netherlands have one integrated steel company (1918-present), one zinc smelter (1892-present), one tin smelter (1928-1994), and two aluminum smelters (1966-present, 1969-present). Trade has always been important because of the presence of the rivers Rhine and Meuse and some canals, which enabled transportation from and to the German and Belgian hinterland, and the presence of colonies, from which present-day Indonesia was the most important one. Geologically, The Dutch state consists of two distinct areas. The western and northern part, in which the major cities and ports are concentrated, has fertile, but low-lying, soils. The eastern and southern part is characterized by higher, but in general infertile, sandy soil and its labor reserves were used for industrial purposes. On this basis, many industries settled there. Agriculture could flourish there only after fertilizers became available.

DATA SEARCH

Following a chain approach, the following data have to be integrated: (1) minerals extraction and agriculture, (2) industrial production, (3) consumption, and (4) waste generation. Not only domestic extraction and production, but also import and export must be considered for both primary products, all major types of industrial products, and waste. In contrast with economically oriented studies, the service industry is not considered and figures on mass instead of value are accounted for. Quantitative data on materials flows in the economy are mainly based on data of the National Bureau of Statistics (CBS) and on censuses performed by the tax authorities. International trade has been documented fairly well over long periods of time and the same holds true for agricultural production. Data on domestic mining and minerals production can be reconstructed, apart from data on peat and bog iron, which are only partly available. A complete set of data on industrial production is not available. A detailed census of 1913 can be used [3], and many data from 1970 are available too. Studies such as those by Brown et al. [2] are useful for the understanding of relationships between mass flows, as well as data derived from statistical materials for traditional processes, such as in [3]. With the help of both studies, mass and energy balances can be derived for a number of basic processes. One must account for the fact that industry not only produces consumer goods but also intermediates.

PRIMARY PRODUCTION

Apart from import and export, domestic production of raw materials is considered in detail. Two types of production are distinguished: Organic materials production, mainly due to agriculture, forestry, fishery and hunting; and minerals extraction, including reclamation of surface minerals, mining, petroleum recovery etc. In agriculture, vegetable and animal production are two distinct domains. Figures are presented in the tables below.

Domestic production of minerals and fossil fuels (kton/yr)				
	1850	**1910**	**1970**	**2005**
Clay (industrial)	1,000	3,000	7,600	3,200
Sand (embankment)	?	?	47,500	43,000
Sand (industrial)	?	?	21,560	17,000
Gravel	?	400	14,380	6,000
Peat	3,500	6,000	1,000	0
Coal	17	1,292	4,334	0
Mineral oil	0	0	1,919	1,492
Natural gas	0	0	25,651	60,313
Bog iron	5	5	1	0
Rock salt (NaCl and $MgCl_2$)	0	0	2,871	5,192
Limestone (industrial)	0	8	2,000	1,580
Limestone (construction)	?	8	1	?

Domestic production in agriculture and forestry (kton/yr)				
	1850	**1910**	**1970**	**2005**
Wheat	129	127	643	1,175
Rye	248	370	172	11
Barley	83	82	334	307
Oat	122	298	201	10
Buckwheat	88	23	0	0
Legumes	82	70	148	25
Vegetables	?	?	521	1,267
Flax (fibers), hemp	7	53	35	37
Oilseed	100	14	30	12
Grass (directly browsed)	?	45,400	22,335	12,425
Hay	?	2,904	3,447	490
Silage	0	0	6,167	9,607
Fodder maize	0	0	302	11,283
Clover etc.	?	?	206	3
Straw	?	?	1,121	?
Potatoes	791	2,179	5,648	6,913
Sugar beet	0	1,367	4,739	5,931
Onions	?	48	619	983
Tomatoes	0	0	358	660
Paprika's	0	0	3	346
Cucumbers	0	0	229	440
Mushrooms	0	0	21	245
Fruit	?	?	565	554
Berries	?	?	58	39
Wood	?	?	1,000	1,000
Fisheries[2]	?	?	?	603

[2] Except from fish farms

Materials that are used in large quantities, but that are neither chemically nor physically transformed, must be considered a distinct category. This refers to materials such as water, embankment sand, and oxygen. Animal production is considered secondary production as animals are nourished with products from agriculture and industry. An exception is the production by hunting and fisheries. Figures on hunting are not available, but in 1850 it was more important than at present. Typical were the duck-decoys that in total cached about 300,000 ducks a year (0,167 kton). Various other primary products too were gathered straight from the environment, such as berries, mushrooms, brushwood, reed, rush, wicker, etc.

IMPORT AND EXPORT

Figures on import and export are available over a wide range of years, see, e.g., [4] and [5]. For commodities that are not domestically produced, the difference between import and export is the available quantity. This holds for many tropical or subtropical agricultural products, stone, and a considerable number of various minerals and industrial products. In case of domestic production, the amount of available products equals:

Available products = domestic production + import – export.

The trade balance for commodities in *currency unit*s has always been negative:
In **1970**, € 22,092×10^6 was imported and € 19,374×10^6 was exported (87,7%)

Expressed in *mass units*, the deficit is still more pronounced:
In **1850**, 1,215 kton was imported and 417 kton was exported (34,3%)
In **1910**, 32,704 kton was imported and 22,489 kton was exported (68,8%)
In **1970**, 118,806 kton was imported and 57,521 kton was exported (48,4%)
In **2005**, cumulative figures are to be determined yet.

Import surplus of some selected commodities (kton/yr)				
	1850	**1910**	**1970**	**2005**
Cereals	85	1,507	3,486	6,433
Oilseeds (incl. Soybeans)	37	271	1,496	3,953
Vegetable oil	**- 3**	16	8	875
Cattle cake	**- 2**	194	1,178	1,056
Mixed feed	0	0	**- 24**	**- 1,908**
Potatoes	**- 8**	**- 167**	**- 852**	**- 1,357**
Vegetables, tomatoes etc.	**- 6**	**- 85**	**- 698**	**- 1,677**
Timber, wood pulp etc.	296	1,144	4,512	5,013
Meat	**- 2**	**- 53**	**- 470**	**- 1,274**
Clay, sand, gravel	9	525	9,394	7,165
Stone, slag	18	707	3,710	7,660
Limestone	0	0	2,286	227
Cement	0	263	1,994	2,090
Iron ore	0	?	5,320	10,308
Raw iron	13	270	171	266
Ferrous scrap	**- 1**	**- 65**	**- 517**	**- 2,110**
Salt	29	102	**- 1,648**	**- 3,751**
Coal	501	6,331	2,441	13,515
Mineral oil	0	103	58,514	51,248
Natural gas	0	0	**- 470**	**- 22,435**

SECONDARY PRODUCTION

Secondary production includes both secondary agricultural production, such as animal production, and industrial production. A complication is here that the output of a specific industry often will be the input of another industry. For instance, the greater part of the produced raw iron is an intermediate product that enters into the steel mills. This provokes double counting when not carefully considered.

No complete set of data on industrial production exists. The best sets of data refer to 1913, see [4], and 1970, but virtually no data on secondary production is available for 2005. However, the quantity of commodities that are not domestically produced can be determined from import and export figures only. For animal production, detailed figures are available.

CONSUMPTION

Figures on final consumption have still to be collected. Some coarse data on human food consumption have been published and other data can be derived from qualitative estimates based on lifestyle analysis and, for more recent years, on sales figures.

ACCUMULATION

Accumulation takes place via temporary storage, which is relevant in case the available storage capacity is comparable with or larger than the yearly consumption. This holds true for, e.g., oil, iron ore, and coal. However, the main contribution is due to the growth of the material infrastructure: the growing amount of buildings, infrastructure and machinery that have a lifespan of decades combined with a large mass. This particularly refers to products made from construction materials such as brick, stone, concrete, asphalt. The increasing quantity of durable commodities such as cars, machinery, furniture etc. must also be accounted for.

BALANCING

Balancing is not always straightforward, because many products have low dry matter content. For milk, e.g., the dry matter content is 13,6%, for grass it is about 20%, for potatoes 25%, for dried peat 30%. Simple operations, such as drying, will considerably affect the mass of a product. Similar considerations can be made with regard to consumption processes. In those processes, combustion and metabolism of living organisms take place, transferring organic materials into mainly carbon dioxide, and water. On the other hand, the mass of manure exceeds that of feed because of the enhanced water content. Chemical reactions, e.g. in cement production, are also affecting the balance.

WASTE

Figures on import and export include those solid wastes that have a market value and that are traded. This refers, e.g., to organic byproducts, scrap metals, old clothes, waste plastics, cullet, wastepaper, etc. Domestic waste production has not been documented very well till in the final decades of the 20th century. Emissions to the surface water and to the atmosphere are not accounted for in this study. A considerable contribution of waste is due to construction and demolition waste, which is rather persistent and is used for paving purposes. Considerable amounts of waste have also been accumulated in waste dumps.

CONCLUSIONS

Although much work still has to be done, some conclusions can already be drawn:

1850

It appears that, in the year 1850, much of the industry was still traditionally organized and based on craftsmanship and home industry rather than on centralized industrial production.

Much of the economy took place out of the domain of the market. Self-subsistence was dominant in the countryside. Both the number of products that were internationally traded and their quantities were modest in most cases. Many of these were organic materials. Other product types included luxury goods. However, coal represented the most massive input although the number of steam engines still was small and much peat still was used for spatial heating. Coal, however, was applied in smithies and in some other processes. Some locomotives were operational as well. Chemicals and complex mechanical products played a minor role in the mass balance. Trade of many agricultural products was restricted by a lack of conservation methods.

1910

Although the industrial revolution based on steam had taken place, the contribution of complex products remained modest. Much coal was imported and a considerable domestic production has started as well. Only a few novel products were present in the statistics, from which fertilizer, beet sugar, and margarine can be mentioned in particular. This and the founding of co-operative agricultural organizations started to boost agricultural production and food processing industry. This reflected itself in the increasing amount of animal feedstuff and fertilizer that was imported. Although the importance of metals was increasing, its significance in the mass balance remained modest. The same was true for mineral oil. There were only 3,000 cars in The Netherlands at that time.

1970

From 1910 on, domestic manufacturing industry has grown tremendously. Process industries also arose on a considerable scale. During World War I, self-sufficiency was aimed for and many basic industries (iron and steel, cement) were erected. Electrification and the massive use of petrol and, subsequently, natural gas as a fuel changed the complete industrial picture. A variety of products were produced, from which a multitude of chemicals, synthetic fibers, plastics, fertilizers, mechanical, electrical and electronic products can be mentioned. Decline, however, already started and many basic industrial sectors were beginning to close down or shifting their activities to low-wage countries. Agricultural production was increasing and particularly factory farming was growing rapidly. In the decades to follow, livestock of pig and poultry would increase dramatically, as well as the import of animal feedstuff.

2005

Flows of imported and exported goods further increased, agricultural productivity further developed, but manufacturing shifted nearly completely to the emerging economies abroad, particularly in South-Eastern Asia and Eastern Europe. Apart from process industries that were still present, the manufacturing industry became mainly restricted to niche-markets and specialized machinery and installations.

REFERENCES

[1] Brunner, P.H. and Rechberger, H., *Material Flow Analysis: Advanced Methods in Resource and Waste Management*. Boca Raton, FL: Lewis Publishers, 2004.

[2] Brown, H.L., Hamel, B.B., and Hedman, B.A., *Energy Analysis of 108 Industrial Processes*. Fairmont Press, 1985.

[3] Centraal Bureau voor de Statistiek (Dutch National Bureau of Statistics), *Statistiek van de Voortbrenging en het Verbruik der Nederlandsche Nijverheid in 1913 en 1916* (Statistics on Industrial Production and Consumpion in The Netherlands in 1913 and 1916)

[4] ibid., *Statistiek van den Handel en de Scheepvaart van het Koningrijk der Nederlanden, over het jaar 1850* (Statistics on Trade and Shipments of the Kingdom of The Netherlands). ’s-Gravenhage: 1851

[5] ibid., *Maandstatistiek van de Buitenlandse Handel, per goederensoort* (Monthly Statistics on International Trade, by commodity), **21**(12) (december 1970). ’s-Gravenhage: 1971

OPTIMAL PRODUCTION CONTROL AND DYNAMIC PRICING STRATEGIES: A JOINT APPROACH

Bin Zhou
College of Business and Public Administration
Kean University, Union, NJ 07083
E-mail bzhou@kean.edu

Hua Zhong
Economics & Business Division
State University of New York, Oneonta
Oneonta, NY 13820, USA
E-mail zhongh@oneonta.edu

ABSTRACT

Production/inventory control and dynamic pricing are important areas in Operations Research and they reflect a wide-range of business practices in Supply Chain systems such as retail, consumer products, and airline industries. Managers encounter challenges from both sides of the Supply Chain and have to make critical decisions such as effective production and/or replenishment policies as well as pricing strategies. Primary research questions include: a) what should be the optimal production/ordering quantity? b) what should be the optimal selling price? In this paper, we study production-pricing models that generate effective solutions to these questions and deliver simple guidelines for management.

Key word: Supply chain; Production and inventory systems; Pricing

INTRODUCTION

Management Science and Operations Research have been areas of significant importance and their research and applications attract substantial attention from both academic researchers and business practitioners for decades. Effective Supply Chain Management (SCM) strategies are the keys of modern organizations in virtually any countries and regions. Production and inventory management, in particular, also accounts for a substantial share of corporate expenditures in these companies, and its cost is expected to become even larger as the requirements and expectations from customers continue to grow. On the other hand, business practices demonstrate that requirements for effective but simple guidelines and principles for managing multiple functional areas increase significantly in the past decade as the complexity of the supply chain system and the need for global optimization continue to grow. However, despite the long-standing theoretical interest in independent production planning policies or dynamic pricing and revenue management strategies, some important issues in the real-life supply chain problems, such as the interaction and coordination among these different areas, are still left untouched in most of the research work.

Both production/inventory and pricing problems are often encountered in real supply chains where companies need to determine production/inventory schedules as well as selling prices for the products offered. The joint operation strategies in these two areas are critical in that if the production/inventory size is beyond that of the demand, appropriate adjustment of the price may help to reduce potential loss; if the production/inventory size is below that of the demand, appropriate adjustment of the price can help to increase potential revenue, and vice versa. This joint problem arises and reflects a wide-range of business practices. For instance, it is a common practice in apparel industry with seasonal products, such as fashion apparel producer and retailer GAP or H&M, who offers promotion and discounts and replenishes its inventory seasonably, i.e. weekly, monthly, or with specified start and end dates. It is more common in retail industry, consumer products (electronic and grocery) industry for repeated ordering and selling activities. For instance, Wal-Mart Stores or Best Buy restocks products from suppliers and promotes products in repeated sales cycles. Accordingly, the interesting questions that researchers and practitioners have dealt with are how many products should be produced/ordered? When should the selling price be adjusted? And what should be the optimal selling price? This research aims to study these problems and propose simple and effective tools and guidelines for both academic research and business practices.

Research on the independent problems of production/inventory control or dynamic pricing is abundant. In the first area, important findings are discovered and widely known by both academic researchers and business practitioners. For the problem with fixed setup costs, the (s, S) type policy is provided in Clark and Scarf [1]. Federgruen and Zipkin [2] studied single-item, multi-period inventory systems with limited production capacity and uncertain demands under both the average-cost and the discounted-cost criteria. Kapuscinski and Tayur [3] analyzed a capacitated production and inventory model with periodic demand.

For systems without fixed setup costs, Zhao and Katehakis [4] investigated a single-item stochastic inventory system with a minimum order quantity in finite and infinite horizons under discounted cost. They showed that the optimal policy is inevitably complex and partially characterized the structure of the policies in certain regions for the multi-period model. Zhou et al. [5] later developed simple and effective heuristics for the single-item problem. The paper showed that this policy significantly outperforms the traditional (s, S) policy under general conditions.

In the area of revenue management, to our knowledge, most work in existing literature is done under the assumptions of continuous-time and finite sales horizon with fixed time periods. Feng and Gallego [6] investigated a two-price, continuous-time yield management model that allows a single price change from a given initial price to either a lower or a higher second price. Bitran and Mondschein [7] focused on the application on the sales of seasonal products. Zhao and Zheng [8] studied a more general continuous-time dynamic airline yield management model and showed that a threshold policy is optimal.

In this article, we propose a model of production and dynamic pricing in the situation where the planning horizon consists of repeating sales cycles, which may represent days, seasons, etc. The demand in each time period of a cycle will have a discrete distribution, which is a function of the selling price and of the time period. The offered price can be selected from a finite set of acceptable prices, and both price increases and decreases are in general possible.

The remainder of this paper is structured as follows. In Section 2, we propose the model and discuss some preliminary conditions. A numerical example is discussed in Section 3 using linear programming. Finally, Section 4 concludes the paper with remarks and future research.

MODEL FORMULATION

We consider a retailer who manages a single item with infinite sales horizon, which includes repeated sales cycles. The retailer produces and replenishes the inventory at the beginning of each sales cycle. We use indices (t, n), where $t = 1, 2, \ldots$ denotes the t^{th} cycle and $n = N, \ldots, 1$ denotes the n^{th} period in cycle t. The retailer can change the offering price from a finite set denoted by $A = \{a_1, \ldots, a_K\}$, and $a_1 > a_K$. Let D_{ak}^n denote the demand when in time period (t, n), the selling price offered is a_k. The demand has a discrete distribution specified by the probabilities $p_{ak}^n(d) = P(D_{ak}^n = d)$, $d = 0,1,\ldots$.We assume a unit stock out penalty cost l is applied if demand cannot be fulfilled from the inventory at any time period. Let β be the discount factor, where $0 < \beta < 1$.

Denote x to be the inventory position at time period n in a sale cycle, and d to be the demand of the period. Given selling price a_k and time period n, the single period reward function of the retailer can be written as

$$r_{ak}(x,n) = R_{ak}(x,n) - L_{ak}(x,n) \tag{1}$$

where

$$R_{ak}(x,n) = a_k\left(\sum_{d=0}^{x} p_{ak}^n(d)d + x\sum_{d=x+1}^{\infty} p_{ak}^n(d)\right) \text{ and } L_{ak}(x,n) = l\left(\sum_{d=x+1}^{\infty} p_{ak}^n(d)(d-x)\right).$$

Further, let $W_{ak}^t(x,n)$ be the expected discounted reward when price a_k is employed at time period (t, n).. Let $V_t(x,n) = \max_{ak\in A} W_{ak}^t(x,n)$ represent the value function. For simplicity, we will use the (t, 0) for time period (t-1, N). Also let $p_{x,y,ak}^n = P(D_{ak}^n = x - y)$ be the transition probability from inventory position x to position y when price a_k is implemented in time period n. Then we obtain the following recursive equations.

$$V_t(x,n) = \max_{ak\in A}\{W_{ak}^t(x,n)\}, \forall t \ge 1 \tag{2}$$

$$V_t(x,0) = V_{t-1}(x_0, N), \forall t \ge 2 \tag{3}$$

Where $W_{ak}^{t}(x,n)=r_{ak}(x,n)+\beta\sum_{y=0}^{x}p_{x,y,ak}^{n}V_t(y,n-1),\ \forall t\geq 1,\ W_{ak}^{t}(x,0)=V_{t-1}(x_0,N),\ \forall t\geq 2$.

Without loss of generality, the salvage value of any unsold unit at the end of the first sales cycle is zero, $V_1(x_0,0,\beta)=W_{ak}^{1}(x,0,\beta)=0,\ \forall x\ ,\ a_k\in A$. Then, the optimality equations under total discounted reward are given as follows:

$$V(x,n)=\max{}_{ak\in A}\left\{r_{ak}(x,n)+\beta\sum_{y=0}^{x}p_{x,y,ak}^{n}V(y,n-1)\right\},\ \forall x,\ n\geq 1, \tag{4}$$

$$V(x,0,\beta)=V(x_0,N,\beta),\ \forall x. \tag{5}$$

Instead of prescribing a particular initial inventory position, we shall determine a quantity that brings maximal expected revenue from the state space I = {0,1,..., M}, where M is the given upper bound due to production or warehouse capacity constraint. We denote x_0^t to be the decision of production/replenishment quantity at the start of sales cycle t, hence the optimality equations that include the production decision can be expressed as follows.

$$W_{ak}^{t}(x,n)=r_{ak}(x,n)+\beta\sum_{y=0}^{x}p_{x,y,ak}^{n}V_t(y,n-1), \tag{6}$$

$$V_t(x,n)=\max{}_{ak\in A}\left\{W_{ak}^{t}(x,n)\right\},\ x\in I,\ a_k\in A,\ n=1,...,N,\ t\geq 1, \tag{7}$$

$$V_t(x,0)=V_{t-1}\left(x_0^{t-1},N\right),\ t\geq 2. \tag{8}$$

Using the above functions, one can solve for the optimal production policy and pricing policy as the same time. We now give sufficient conditions under which the optimal policy possesses the mark down structure in remaining inventory.

Condition 3.1 *In any time period n, given selling prices a_{k-1} and a_k , with $a_{k-1} > a_k$ and inventory position x, the following conditions are true*

$$\frac{a_k+l}{a_{k+1}+l}\geq \max\left\{\frac{P\left(D_{ak-1}^{n}\geq x\right)}{P\left(D_{ak}^{n}\geq x\right)}\right\} \tag{9}$$

From the practical point of view, the condition imposed on the demand probabilities is consistent with the real-world situation that demands at higher price are generally stochastically less than those at lower price, i.e. $D_{ak-1}^{n}\leq D_{ak}^{n}$.

NUMERICAL EXAMPLE

We can compute the optimal pricing using a linear programming model. We simplify the notation and let $a_k = k$ and re-write the transition probability as $p_{x,y,k}^{n}$. Let $Y_{x,k}$ denote the long-term fraction of time at state x when taking action k. We present the following linear programming model to compute the optimal pricing policies and the value functions of the retailer.

$$
\begin{aligned}
\max \quad & \sum_{x\in X}\sum_{k\in K} r_{x,k}\cdot Y_{x,k} \\
s.t. \quad & \sum_{x\in X}\sum_{k\in K} Y_{x,k} = 1 \\
& \sum_{k\notin K} Y_{y,k} = \sum_{x\in X}\sum_{k\in K} Y_{x,k}\cdot p^{n}_{x,y,k},\ y = 0,\ldots,X \\
& Y_{x,k} \geq 0
\end{aligned}
\tag{10}
$$

where $r_{x,k}$ is the single period reward function.

Table 1. Optimal Pricing Policy with Mark Down Structure in Remaining Inventory. (Example 1)

Inventory position	Day 6	Day 5	Day 4	Day 3	Day 2	Day 1
1	-	H	H	H	H	M
2	-	H	H	H	H	M
3	-	H	H	H	H	M
4	-	H	H	H	M	M
5	-	H	H	H	M	M
6	-	H	H	H	M	M
7	-	H	H	H	M	L
8	-	H	H	M	M	L
9	-	H	H	M	M	L
10	H	H	H	M	M	L

We now give a simulated example. Assume the unit penalty cost l = 10. Suppose the retailer is selling a product to the market. The sales cycle consists of 6 days, i.e., N = 6, which starts in day 6 and ends in day 1. Initial inventory position is 10. Demand is discrete from 0 to 10. The product can be sold at high price a_1 = \$50, medium price a_2 = \$48, or low price a_3 = \$44. For the first three days, n = {4, 5, 6}, the probabilities for no demand are $p^n_{a1}(0) = 0.55$ at high price, $p^n_{a2}(0) = 0.35$ at medium price, and $p^n_{a3}(0) = 0.2$ at low price. If there is demand, $p^n_{a1}(1,\ldots,10) = 0.045$, $p^n_{a2}(1,\ldots,10) = 0.065$, and $p^n_{a3}(1,\ldots,10) = 0.08$ at three prices respectively. For the last three days, n = {1, 2, 3}, the probabilities for no demand are $p^n_{a1}(0) = 0.5$, $p^n_{a2}(0) = 0.25$, and $p^n_{a3}(0) = 0.05$. And the probabilities for non-zero demand for the products are, $p^n_{a1}(1,\ldots,10) = 0.05$, $p^n_{a2}(1,\ldots,10) = 0.075$, and $p^n_{a3}(1,\ldots,10) = 0.095$. Table 1 presents our findings. For emphasis in the table, we use the notation $H = a_1$ = \$50, $M = a_2$ = \$48, and $L = a_3$ = \$44.

The results in Table 1 illustrate the mark down structure of the optimal policy. At a give time period, the optimal policy decreases as the remaining inventory position x increases. For instance, in day 2, if there are 4 or more inventory left, the optimal price should be adjusted to the medium price a_2 = \$48. If there are less than 4 units, the retailer should continue to sell the product at the highest price a_1 = \$50, to achieve maximal expected revenues.

CONCLUSION AND FUTURE RESEARCH

In this article, we have discussed a joint production/inventory and pricing model in which demand is a random variable that has a discrete distribution, and is dependent on both selling price and time period. The product is sold over an infinite horizon, where each sales cycle contains a finite number of time periods. Price adjustment can be executed in both directions from among a prescribed finite set of allowable prices. We provide simple conditions on the demand probabilities and costs for the optimal production and pricing policies. We are considering several directions in which interesting research questions still remain. One direction of future research is that price reversibility is costly, thus price adjustment is only allowed in the downward direction. In other words, once the price is changed from a higher price to a lower one, it cannot be switched back. In such a case, optimal policy has a mark down structure in time period.

REFERENCES

[1] Clark, A. & S. Scarf. Optimal policies for a multi-echelon inventory problem. *Management Science*, 1960, 475-490.

[2] Federgruen, A., & P. Zipkin. An inventory model with limited production capacity and uncertain demands I: the average cost criterion. *Mathematics of Operations Research*, 1986, 11, 193-207.

[3] Kapuscinski, R., & S. Tayur. A capacitated production-inventory model with periodic demand. *Operations Research*, 1998, 46, 899-911.

[4] Zhao, Y., & M.N. Katehakis. On the Structure of Optimal Ordering Policy for Stochastic Inventory Systems with Minimum Order Quantity. *Probability in the Engineering and Informational Sciences*, 2006, 20, 257-270

[5] Zhou, B., Y. Zhao, & M. N. Katehakis. “Effective control policies for stochastic inventory systems with a minimum order quantity and linear costs”, *International Journal of Production Economics*, 2007, 106, 523-531.

[6] Feng, Y., & G. Gallego. Optimal stopping times for end of season sales and optimal stopping times for promotional fares. *Management Science*, 1995, 41, 1371-1391.

[7] Bitran, G. R., & S. V. Mondschein. Periodic pricing of seasonal products in retailing. *Management Science*, 1997, 43, 64-79.

[8] Zhao, W., & Y. S. Zheng. Optimal dynamic pricing for perishable assets with nonhomogeneous demand. *Management Science*, 2000, 46, 375-388.

USE OF AHP AND TAGUCHI LOSS FUNCTIONS FOR SUPPLIER SELECTION

Sharon M. Ordoobadi
University of Massachusetts-Dartmouth
285 Old Westport Road, N. Dartmouth, MA. 02748
sordoobadi@umassd.edu
508-999-8767

ABSTRACT

Supplier selection is a multi-criteria decision making process encompassing various tangible and intangible factors. Both risks and benefits of using a vendor in supply chain are identified for inclusion in the evaluation process. A hybrid approach that applies to both quantitative and qualitative factors is used in the development of the model. Taguchi loss functions are used to measure performance of each supplier with respect to the risks and benefits. Analytic Hierarchy Process (AHP) is used to determine the relative importance of these factors to the decision maker. The weighted loss scores are then calculated for each supplier by using the relative importance as the weights. The composite weighted loss scores are used for ranking of the suppliers. The supplier with the smallest loss score is recommended for selection.

INTRODUCTION

With growth of the world economy and global marketplace the importance of outsourcing is more than ever. Many firms are engaged in outsourcing a number of their selected activities in order to take advantage of the external expertise. Although outsourcing provides the firms with many benefits, if the right activity is not chosen for outsourcing or inappropriate supplier is chosen to perform the activity the negative impacts on the firm could be drastic. Thus it is crucial for the decision makers to analyze the impacts of their outsourcing policy on their organizations. The aim of the current research is to provide a tool for performing such an analysis and make this tool accessible to the decision makers.

To allow a thorough analysis of an outsourcing decision the tool should provide ample opportunity for the decision maker to: first, be aware of pros and cons of outsourcing, second, be able to compare these positive and negative impacts in a systematic and scientific way, and finally be able to rank potential suppliers for performing the outsourcing function. The proposed model is developed in order to serve these purposes. Risk and benefit categories associated with outsourcing are identified. The subjective judgments of the decision makers on the importance of these factors are elicited through the use of AHP. The pair-wise comparisons are used to determine the priority weights of the risks and benefits. The performance of each supplier with respect to the risk and benefit categories are measured by a common value, Taguchi's loss score. The weighted loss scores are determined by combining the loss scores and the weights from the AHP. The composite weighted loss scores are calculated for each potential supplier. The suppliers are then ranked according to their composite weighted scores.

The rest of the paper is organized as follows: The background and overview of the literature is provided in section 2. The development of the proposed model is detailed in section 3. The paper concludes with summary and conclusions in section 4.

II. LITERATURE REVIEW

In today's economy outsourcing has become a norm rather than an exception and thus it is very important to identify the appropriate supplier to outsource to. Supplier selection is a multi-criteria decision making process encompassing various tangible and intangible factors. These factors however do not have the same weight in the decision making process. The level of importance of these criteria varies by the decision maker and the situation at hand. As a result most of the supplier selection models have been developed with the capability of addressing this issue. A brief overview of the literature is provided below. Numerous supplier selection models are uncovered in literature. These methods fall into one the following categories:

Categorical methods (Timmerman, 1986); Supplier's performance on all relevant criteria are categorized as "positive", "neutral", and "negative". The supplier whose performance receives the most "positive" values is the best.

Data envelopment analysis (DEA) (Weber & Desai, 1996; Weber *et al*. 1998; Liu *et al*. 2000); The efficiency of each supplier is calculated as the ratio of weighted sum of its output (the performance of the supplier) to the weighted sum of its input (the cost of using the supplier)

Cluster Analysis (CA) (Holt, 1998); Uses a classification algorithm to group a number of suppliers into clusters based on a set of numerical attribute scores.

Linear Weighting models (Timmerman, 1986)*;* Weights are assigned to the criteria based on their importance. Ratings on the criteria are multiplied by their weights and summed to obtain a single score for each supplier. The supplier with the highest overall ratings is selected. Variations to the basic linear weighting models are proposed by various researchers. Analytical Hierarchy Process (AHP) is used by others (Nydick & Hill 1992; Barbarosoglu & Yazagac 1997; Masella & Rangone 2000, Sarkis & Talluri, 2000).

Mathematical programming models (Chaudhry *et al.* 1993; Weber & Desai 1998; Ghoudsypour & O'Brien 1998; Weber *et al.* 2000); The supplier selection problem is formulated in terms of an objective function. The appropriate supplier is identified by finding the optimal solution.

In most of the above models the appropriate supplier is selected based on the selection criteria. Very few models have included the risks and benefits of outsourcing in the evaluation process. Even fewer attempts have been made to provide a systematic approach for quantifying the intangible risks and benefits. To fill such a gap a hybrid approach is proposed here that could address both issues; the inclusion of risks and benefits which are both qualitative and quantitative in nature as well as a systematic approach to quantification of intangibles. AHP methodology is used to include decision makers' judgments regarding the risks and benefits. Taguchi loss functions are used as a mechanism to quantitatively represent the suppliers' scores by measuring their performances with respect to these risks and benefits. By employing AHP and Taguchi a superior approach to supplier selection is provided. The development of the proposed model is detailed in the next section.

III. DEVELOPMENT OF THE MODEL

The premise of the proposed model is that the outsourcing decision is justified both strategically and financially. That is the outsourcing decision does not compromise the firm's competitive position, and is in line with strategic goals of the organization. In addition, the firm can take advantage of the monetary benefits by engaging in an outsourcing decision. The purpose of the proposed model is to identify the appropriate supplier to carry out the outsourcing function.
The model is developed by completing the following steps:

- Identifying the risk and benefit categories associated with outsourcing.
- Determining the relative importance of each risk and benefit category.
- Determining the performance of each supplier with respect to the risk and benefit categories.
- Developing a mechanism to quantitatively measure the suppliers' performance with respect to all the pertinent risks and benefits.
- Determining the rankings of the potential supplier and selecting the appropriate supplier.

Each of the above steps is detailed in the following sections.

Identifying risks and benefits of outsourcing

Any sourcing decision has certain benefits and risks that should be included in the evaluation process. A careful review of the literature was conducted to uncover the categories of risk and benefit associated with an outsourcing decision. A list of possible benefits and risks that could materialize from a sourcing decision is compiled from the available literature on the subject.
Some benefits offered by outsourcing cited in the literature (Chalos 1995; McCarthy 1996; Parasuramann *et al.* 1988) are:

- Higher level of flexibility, with less restriction from the rules existing in the company.
- Increased responsiveness to customers' needs.
- Providing special services to the customers through outsourcing without the need to hire special skill workers.
- Liability and risk reduction.
- Reduction of capital investment and labor requirements.
- Access to the innovations and developments of more specialized suppliers.
- Reduction of costs due to the supplier's economies of scale.
- Greater focus of resources on high value-added activities and core business.
- Assurance (competence, courtesy, credibility, security).
- Responsiveness.
- Empathy (communication, access, understanding).

Some of the risks associated with outsourcing cited in the literature (Friedman 1991; Raistrick 1993) are:

- Lack of control on quality of the product/service provided by the suppliers.
- Inability to meet fluctuations in demand for the product/service that has been outsourced.
- Loss of control over suppliers.

- Negative impact on employees' morale.
- Loss of critical skills or developing the wrong skills.
- Loss of cross-functional skills.

Of-course not all the risks and benefits in the above lists are pertinent to every situation and every decision maker. The list of the benefits and risks is narrowed down to include only the relevant categories. This is accomplished through a calibration procedure where the decision maker is asked to identify the benefits/risks pertinent to his/her situation. In addition, the selected categories have different level of importance to the decision maker. Thus, there is a need to determine the relative importance of each risk and benefit category. This is accomplished by completing the second phase of the model development explained in the following section.

Determining the relative importance of each risk and benefit category

The importance of each risk and benefit is subjective and varies by the decision maker. To capture the subjective judgments of the decision maker the AHP methodology is used. AHP introduced by (Saaty, 1995) is widely used for solving multi-attribute decision making problems. In the current research the AHP methodology is used to determine the relative importance of various factors considered in the evaluation of the potential suppliers. This is accomplished by performing several pair-wise comparisons where the decision maker states the importance of one criterion over the rest of the criteria on the scale of (1-9). Where 1 means least preferred and 9 means most preferred and numbers between the two extreme shows moderate importance.

The pair-wise comparisons are performed for benefits and local priorities are calculated using the approach recommended by (Saaty, 1995). These priorities represent the relative importance of the benefit categories. The local priorities for the risk categories are calculated in the same manner. These priorities are then used as weights in calculation of suppliers' weighted loss scores. Once the importance of each risk and benefit category is determined there is a need to find out how each potential supplier perform with respect to these risk and benefit categories. This is achieved by completing the third phase of the model development as explained in the next section.

Determining the performances of suppliers with respect to the risk and benefit categories

Suppliers' performances vary with respect to the pertinent risk and benefit categories. More importantly two decision makers might have different perception of the performance of the same supplier with respect to the exact same risk or benefit category. Thus, inclusion of the subjective judgment of the decision maker in the evaluation process is crucial. To achieve this, an elicitation procedure is performed to solicit the decision maker's perceptions of suppliers' performances. The decision maker often base such judgments on the historical data, reputation of the supplier, the specifics of the situation at hand, etc. Once pertinent risk and benefit categories are selected and potential suppliers are identified the elicitation process is performed.

At this point the decision maker has identified all the pertinent risks and benefits, the relative importance of each are determined through application of AHP methodology. Furthermore, the decision maker's perceptions of suppliers' performance with respect to these risk and benefit categories have been elicited. However, to allow ranking of the potential suppliers the results of the analysis performed so far should be combined and subjective perceptions ought to be quantified to come up with a single performance score for each supplier. The Taguchi loss function was considered as a means to accomplish this task as explained in the next phase of the model development.

Developing a mechanism to quantitatively measure suppliers' performances with respect to risk and benefit categories.

Generally, three types of loss functions are used to calculate Taguchi loss (Taguchi & Hsiang 1989; Besterfield *et al.* 2003). First, two sided loss function where nominal value is the target and deviation from either side of the target is allowed as long as it remains within the specification limits. The second and third types of loss functions are one-sided functions where deviations from the target are allowed only in one direction. These loss functions are referred to as "larger-is better" and "smaller-is-better".

For the purpose of this study, one-sided minimum specification limit Taguchi loss function (larger-is-better) has been used to quantify the impact of outsourcing benefits. The rationale is that the more outsourcing benefits received from a supplier compared to performing the function in-house is better. 100% possibility of receiving an outsourcing benefit over the in-house performance is the target value. The deviation from this target value is possible only in one direction and the magnitude of allowable deviation is set by the decision maker. Such a loss function is formulated as $L(X) = k\ (1/X^2)$ where X is a vector representing desired specification limits set by the decision maker for the pertinent benefit categories. Thus, if the number of relevant benefit categories is denoted by n then X: x_i : i=1,2,…,n. L(X) is the loss for specific value of X, and k is the loss coefficient whose value depends on the specification limits set by the decision maker.

Using such a one-sided loss function, 100% Taguchi loss occurs at the lowest specification limit. For instance, the decision maker might set the lowest specification limit for delivery of a benefit category, compared to performing the function in-house, at 70%. Then the Taguchi loss is 100% for a supplier that performs at this level. The value of the loss coefficient k is calculated as $100^* (0.70)^2 = 49$, and the loss function for this particular benefit category is identified as $L(X) = 49(1/X^2)$. Thus, the loss score for suppliers whose performance meet the specification limit can be calculated using this loss function. Appropriate loss functions for all the benefit categories are determined in similar manner. Supplier's loss scores are then calculated using these loss functions. As a result each supplier will end up with several separate loss scores for all the pertinent benefit categories. However, a single value is desirable to allow the comparison of the performance of the potential suppliers. To achieve this, for each supplier a weighted loss score for all the benefit categories is calculated. The weights used in the calculation are the relative importance of each benefit category that has already been determined through AHP methodology.

The one-sided maximum specification limit loss function (smaller-is-better) is used to quantify the impact of the risk categories on an outsourcing decision. The target value is zero, thus the loss is zero when delivery of a risk category by a supplier is unlikely. The loss is 100% when the risk is at its upper specification limit. Such a loss function is formulated as $L(Y) = kY^2$; Where vector Y represents the specification limits set by the decision maker for the relevant risk categories. If m denotes the number of these risk categories, then Y: y_i: i=1, 2,…,m. L(Y) is the loss for specific value of Y and k is the loss coefficient.

Using decision maker's specification limits and the above loss function, the loss coefficient k can be calculated. For instance, a decision maker might set the specification limit of 5 days for responding to fluctuations in demand. Then the Taguchi loss for a supplier who can respond immediately is zero while the loss for a supplier who responds in 5 days is 100%. The loss coefficient is calculated as $k = 100 / (5)^2 = 4$, and the loss function for this risk category is determined as $L(Y) = 4Y^2$. Thus, suppliers' loss scores for those who can respond faster than the limit of 5 days can be calculated using this loss function. The loss functions for the rest of the risk categories are determined in a similar fashion. The supplier's loss score for these risk categories are calculated using the appropriate loss functions. In order to come up with a single loss score for all the risk categories, the supplier's individual loss scores and the relative importance of these risk categories are used to calculate a weighted loss score.

To select the appropriate supplier, potential suppliers need to be ranked based on their composite loss scores that combine the weighted loss scores for benefits and risks. The calculation of the composite loss scores and ranking of the suppliers are performed in the final stage of the model as detailed in the following section.

Determining rankings of the suppliers

At this point each supplier has received a weighted loss score for all pertinent benefit categories as well as a weighted loss score for all relevant risk categories. However, to compare performances of the potential suppliers, a single loss score for each supplier is desirable. To accomplish this task, the average loss score for each supplier is determined by calculating the average of the weighted loss scores for benefit and cost categories that have already been calculated.

The suppliers are then ranked based on their average loss scores. In this study, a final unified loss score was obtained by using the average of the loss scores for benefits and risks. Of-course, the calculation of the composite loss score can vary by the company and/or the decision maker. The proposed model can be used to consider various factors for ranking of the suppliers.

IV. CONCLUSIONS

A decision model is developed to help decision makers with selection of the appropriate supplier for the outsourcing purposes. Although outsourcing provides certain benefits for the company it carries with it several risks as well. The proposed decision model is an attempt to consider all aspects of an outsourcing decision in the evaluation process. To accomplish this task first all benefit and risk categories associated with outsourcing are identified. The importance of each category along with decision maker's perception of supplier's performance with respect to these categories is elicited. AHP is used to determine the relative importance of each category and Taguchi loss functions are applied to quantitatively measure the supplier's performance. Individual loss scores for each benefit/cost category are calculated. The weighted and composite loss scores are then determined and used for ranking of the potential suppliers. The supplier with the lowest composite loss score is chosen to perform the outsourcing function.

(References are available upon request from Sharon Ordoobadi)

EVALUATING MULTISTAGE SUPPLY CHAIN EFFICIENCY WITH DEA MODELS

Ke Wang, Beihang University, University of Illinois, (217) 244-1074, kewang09@uiuc.edu
Fajie Wei, Beihang University, (8610) 8231-6180, weifajie@buaa.edu.cn

ABSTRACT

In order to evaluate the multistage supply chain efficiency, an appropriate performance evaluation system is importantly required. A representative multistage supply chain has three members composing a supplier-manufacturer-retailer structure, and has intermediate measures connecting these supply chain members. Existing data envelopment analysis (DEA) models have difficulties in measuring these kinds of supply chain efficiencies directly. In our current study, we develop several DEA based models for characterizing and measuring these multistage supply chain efficiencies, and considering the intermediate measures. We illustrated the models in a three-stage supply chain context which could represent different relationships between the supplier, manufacturer and retailer when they are treated as different supply chain structures: 1) the non-cooperative, 2) the partial-cooperative, and 3) the cooperative supply chain structure. We also demonstrate these models with a numerical example.
Keywords: Data envelopment analysis, Efficiency, Evaluation, Multistage, Supply chain

INTRODUCTION

Supply chain performance evaluation is a basic work for an organization to promote it supply chain efficiency. In order to evaluate the multistage supply chain efficiency, an appropriate performance evaluation system is importantly required. Data envelopment analysis (DEA)[1], first proposed by Charnes et al., is a linear programming based methodology for evaluating the relative efficiency of each member of a set of organizational units, which are called decision making units (DMUs). DEA evaluates the efficiency of each DMU relative to an empirical production possibility frontier[2]. The traditional standard DEA model does not consider the internal structure of the DMUs, i.e., it treats each DMU as a black box by considering only the input consumed and the output produced by each DMU[3]. This DEA approach has difficulties in measuring supply chain efficiencies, which include the supply chain's efficiency and it members' efficiencies, because it neither provides insight of the internal operations of each DMU nor makes locations of efficiency or inefficiency to the multiple sub stages in each DMU. Thus, the standard DEA approach should be improved to emphasize the stages of the supply chain production process, i.e., the complicated series supply chain production process should be divided into sub process, in that the intermediate products are considered. Especially some of the intermediate products are outputs from a sub process on the one hand and are inputs to another sub process on the other hand. In some researches, the complication production process is composed of two sub processes connected in series. Seiford and Zhu consider a production process of commercial bank as two stages of profitability and marketability[4]. The inputs of the first stage denote the inputs of the bank production process, and the outputs of the second stage denote the outputs of the bank production process. The efficiencies of the first stage, second stage, and the bank whole production process are calculated through three independent DEA models. Whit the similar idea, Zhu analyzes the efficiencies of the 500 fortune companies[5]. Fare and Grosskopf develop a network DEA approach to model general multiple stage processes with intermediate inputs and outputs[6]. Their network model has a general structure which allows to be applied to a variety of situations include the intermediate products, the allocation of fixed factors, and certain dynamic systems. Sexton and Lewis show in their paper how to use DEA approach to model two stage DMUs and apply their model to major league baseball[7]. They demonstrate the model's advantages over standard one stage DEA as: 1) their model could detect the inefficiency that may undetected by the one stage DEA model; 2) their model could provide greater managerial insight into the locations of inefficiency within the whole system. The two stage sub production process concept has also been applied to evaluate the performance of supply chain. Chen and Zhu proposed a DEA framework which considers a two stage process as efficient when each stage is efficient[8]. In the two stage DEA model of Kao and Hwang, the efficiency of the overall process is the product of the efficiencies of the two stages, and the multipliers on the intermediate measures are the same for the two stages[9]. The assumption of same multipliers also links the two stages. Chen et al. point out the limitation of the Kao and Hwang model, and develop an additive two stage DEA approach in which the overall efficiency is expressed as a weighted sum of the efficiencies of the individual stages[10]. Liang et al. develop several DEA based approaches for characterizing and evaluating supply chain and its member efficiencies, when the intermediate measures are considered in the performance evaluation[11]. The two stages of the supply chain in their models are illustrated as a seller-buyer structure, and the relationship between them is treated as a leader-follower and a cooperative relationship.

In our current research, we consider a representative multistage supply chain which has three members composing a supplier-manufacturer-retailer structure, and has intermediate measures connecting these supply chain members. Expending the idea for two stage DMU efficiency measuring in Liang et al.'s paper, and considering not only intermediate inputs and outputs but normal inputs and outputs in Fare and Grosskopf's paper, we develop several DEA based models for characterizing and measuring these multistage supply chain efficiencies. And we illustrated the models in a three stage supply chain context which could represent different relationships between the supplier, manufacturer and retailer when they are treated as different chain structures: 1) non-cooperative, 2) partial-cooperative, and 3) cooperative supply chain structure.

THE NON-COOPERATIVE MULTISTAGE SUPPLY CHAIN MODELS

The multistage supply chain is described in Figure 1, in which the supplier, manufacturer and retailer are denoted by A, B and C. The whole supply chain is considered as a DMU, where X_A, X_B and X_C are the (normal) input vectors of A, B and C respectively, and Y_{A2}, Y_{B2}, and Y_C are the (normal) output vectors of A, B and C, respectively. Y_{A1} is the output vector of A and also the input vector of B, then Y_{B1} is the output vector of B and also the input vector of C. Therefore, Y_{A1} and Y_{B1} are considered as the intermediate measures which link the supply chain members.

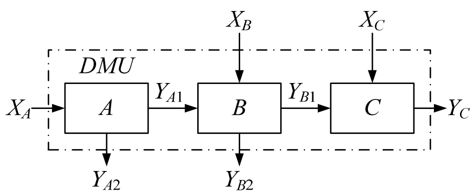


Figure 1. Three stage supply chain structure

Suppose there are n supply chains (DMUs) as above. The CCR DEA model for the supply chain overall efficiency measure is model (1).

$$\max \frac{\sum u_r^{A_2} y_{rj_0}^{A_2} + \sum u_r^{B_2} y_{rj_0}^{B_2} + \sum u_r^{C} y_{rj_0}^{C}}{\sum v_i^{A} x_{ij_0}^{A} + \sum v_i^{B} x_{ij_0}^{B} + \sum v_i^{C} x_{ij_0}^{C}}, \frac{\sum u_r^{A_2} y_{rj}^{A_2} + \sum u_r^{B_2} y_{rj}^{B_2} + \sum u_r^{C} y_{rj}^{C}}{\sum v_i^{A} x_{ij}^{A} + \sum v_i^{B} x_{ij}^{B} + \sum v_i^{C} x_{ij}^{C}} \le 1, u_r^{A_2}, v_i^{A}, u_r^{B_2}, v_i^{B}, u_r^{C}, v_i^{C} \ge 0,\ j = 1,...,n\,. \quad (1)$$

Model (1) only considers the inputs and outputs of the whole supply chain, but misses the intermediate measures which link the supply chain members. Model (1) also could not indicate the individual efficiency of each supply chain member. In order to evaluate the supply chain performance as well as its members' performance, and consider the relationship between the supplier, manufacturer and retailer, we first propose the non-cooperative models, in which the supplier is first evaluated, the manufacturer is second evaluated using the information related to supplier's efficiency, and the retailer is third evaluated using both of the supplier's and manufacturer's information about efficiencies. Firstly, the efficiency of the supplier is measured in model (2).

$$\max \frac{\sum u_r^{A_1} y_{rj_0}^{A_1} + \sum u_r^{A_2} y_{rj_0}^{A_2}}{\sum v_i^{A} x_{ij_0}^{A}} = \theta_A, \text{ s.t. } \frac{\sum u_r^{A_1} y_{rj}^{A_1} + \sum u_r^{A_2} y_{rj}^{A_2}}{\sum v_i^{A} x_{ij}^{A}} \le 1,\ \ u_r^{A_1}, u_r^{A_2}, v_i^{A} \ge 0,\ j = 1,...,n\,. \quad (2)$$

The optimal value θ_A^* is the supplier's efficiency. Then, the efficiency of the manufacturer is measured in model (3).

$$\max \frac{\sum u_r^{B_1} y_{rj_0}^{B_1} + \sum u_r^{B_2} y_{rj_0}^{B_2}}{\sum v_i^{B} x_{ij_0}^{B} + \sum u_r^{A_1} y_{rj_0}^{A_1}} = \theta_{AB}, \qquad \text{s.t.} \quad (3)$$

$$\frac{\sum u_r^{B_1} y_{rj}^{B_1} + \sum u_r^{B_2} y_{rj}^{B_2}}{\sum v_i^{B} x_{ij}^{B} + \sum u_r^{A_1} y_{rj}^{A_1}} \le 1, \frac{\sum u_r^{A_1} y_{rj_0}^{A_1} + \sum u_r^{A_2} y_{rj_0}^{A_2}}{\sum v_i^{A} x_{ij_0}^{A}} = \theta_A^*, \frac{\sum u_r^{A_1} y_{rj}^{A_1} + \sum u_r^{A_2} y_{rj}^{A_2}}{\sum v_i^{A} x_{ij}^{A}} \le 1, u_r^{A_1}, u_r^{A_2}, v_i^{A}, u_r^{B_1}, u_r^{B_2}, v_i^{B} \ge 0,\ j = 1,...,n\,.$$

In model (3), the manufacturer's efficiency is measured under the condition that the supplier's efficiency remains at θ_A^*. The optimal value θ_{AB}^* is the manufacturer's efficiency when the supplier has first achieved its best performance. At last, the efficiency of the retailer is measured in model (4).

$$\max \frac{\sum u_r^{C} y_{rj_0}^{C}}{\sum v_i^{C} x_{ij_0}^{C} + \sum u_r^{B_1} y_{rj_0}^{B_1}} = \theta_{ABC},$$

$$\text{s.t. } \frac{\sum u_r^{C} y_{rj}^{C}}{\sum v_i^{C} x_{ij}^{C} + \sum u_r^{B_1} y_{rj}^{B_1}} \le 1, \ \frac{\sum u_r^{B_1} y_{rj_0}^{B_1} + \sum u_r^{B_2} y_{rj_0}^{B_2}}{\sum v_i^{B} x_{ij_0}^{B} + \sum u_r^{A_1} y_{rj_0}^{A_1}} = \theta_{AB}^*, \ \frac{\sum u_r^{B_1} y_{rj}^{B_1} + \sum u_r^{B_2} y_{rj}^{B_2}}{\sum v_i^{B} x_{ij}^{B} + \sum u_r^{A_1} y_{rj}^{A_1}} \le 1, \quad (4)$$

$$\frac{\sum u_r^{A_1} y_{rj_0}^{A_1} + \sum u_r^{A_2} y_{rj_0}^{A_2}}{\sum v_i^{A} x_{ij_0}^{A}} = \theta_A^*, \ \frac{\sum u_r^{A_1} y_{rj}^{A_1} + \sum u_r^{A_2} y_{rj}^{A_2}}{\sum v_i^{A} x_{ij}^{A}} \le 1,\ \ u_r^{A_1}, u_r^{A_2}, v_i^{A}, u_r^{B_1}, u_r^{B_2}, v_i^{B}, u_r^{C}, v_i^{C} \ge 0,\ j = 1,...,n\,.$$

Under the condition that the supplier has first achieved its beat performance, and the manufacturer's efficiency remains at θ_{AB}^*, model (4) gives the retailer's efficiency θ_{ABC}^* as the optimal value. Then, the efficiency of the whole supply chain could be defined as the arithmetic mean of the supply chain members' efficiencies in (5).

$$E_{ABC} = w_1\theta_A^* + w_2\theta_{AB}^* + w_3\theta_{ABC}^*, \text{ and } \ w_1 + w_2 + w_3 = 1 \quad (5)$$

where w_1, w_2 and w_3 are decision maker specified weights of supplier, manufacturer and retailer, which could reflect their importance in the supply chain, or their power to influence the supply chain. We name the model structure above "forward non-cooperative structure", because the evaluation order is same with the goods flow sequence in a supply chain. The evaluation order also could be reversed, and we call it "backward non-cooperative structure".

THE PARTIAL-COOPERATIVE MULTISTAGE SUPPLY CHAIN MODELS

In this section, we propose the partial-cooperative models, in which an alliance is constructed between supplier and manufacturer, or manufacturer and retailer. The alliance is first evaluated, and then with the alliance's information about efficiency which is defined as the arithmetic mean of the alliance members' efficiencies, the supplier or retailer is evaluated. Suppose supplier and manufacturer are allied and seek to simultaneously maximize both of their efficiencies. The joint efficiency of the alliance is measured in model (6).

$$\max\left[w_A \frac{\sum u_r^{A_1} y_{rj_0}^{A_1} + \sum u_r^{A_2} y_{rj_0}^{A_2}}{\sum v_i^A x_{ij_0}^A} + w_B \frac{\sum u_r^{B_1} y_{rj_0}^{B_1} + \sum u_r^{B_2} y_{rj_0}^{B_2}}{\sum v_i^B x_{ij_0}^B + \sum u_r^{A_1} y_{rj_0}^{A_1}} \right] = \theta_{(AB)}, \tag{6}$$

$$\text{s.t.} \frac{\sum u_r^{A_1} y_{rj}^{A_1} + \sum u_r^{A_2} y_{rj}^{A_2}}{\sum v_i^A x_{ij}^A} \le 1, \frac{\sum u_r^{B_1} y_{rj}^{B_1} + \sum u_r^{B_2} y_{rj}^{B_2}}{\sum v_i^B x_{ij}^B + \sum u_r^{A_1} y_{rj}^{A_1}} \le 1, w_A + w_B = 1, u_r^{A_1}, u_r^{A_2}, v_i^A, u_r^{B_1}, u_r^{B_2}, v_i^B \ge 0, j = 1, ..., n.$$

where w_A and w_B are decision maker specified weights of supplier and manufacturer. Chen et al. point out that model (6) cannot be turned into a linear program using the usual C-C transformation[10]. To solve the problem, they argue that the weights should reflect the "size" of each DMU sub stage, and one reasonable presentation of the "size" is the portion of the total resources devoted to each stage. Based upon their weight choice approach to convert model (6) into a linear program, w_A and w_B could be defined in (7), and the objective function of model (6) could be transformed into (8).

$$w_A = \frac{v_i^A x_{ij_0}^A}{v_i^A x_{ij_0}^A + v_i^B x_{ij_0}^B + u_r^{A_1} y_{rj_0}^{A_1}}, w_B = \frac{v_i^B x_{ij_0}^B + u_r^{A_1} y_{rj_0}^{A_1}}{v_i^A x_{ij_0}^A + v_i^B x_{ij_0}^B + u_r^{A_1} y_{rj_0}^{A_1}}. \tag{7}$$

$$\frac{\sum u_r^{A_1} y_{rj_0}^{A_1} + \sum u_r^{A_2} y_{rj_0}^{A_2} + \sum u_r^{B_1} y_{rj_0}^{B_1} + \sum u_r^{B_2} y_{rj_0}^{B_2}}{\sum v_i^A x_{ij_0}^A + \sum v_i^B x_{ij_0}^B + \sum u_r^{A_1} y_{rj_0}^{A_1}}. \tag{8}$$

The optimal value $\theta_{(AB)}^*$ is the alliance's efficiency, and w_A^* and w_B^* represent the optimal weights obtained from model (6) by the way of (7). Then, the efficiency of the retailer is measured in model (9).

$$\max \frac{\sum u_r^C y_{rj_0}^C}{\sum v_i^C x_{ij_0}^C + \sum u_r^{B_1} y_{rj_0}^{B_1}} = \theta_{(AB)C},$$

$$\text{s.t.} \frac{\sum u_r^C y_{rj}^C}{\sum v_i^C x_{ij}^C + \sum u_r^{B_1} y_{rj}^{B_1}} \le 1, \quad w_A + w_B = 1, \quad \frac{\sum u_r^{A_1} y_{rj}^{A_1} + \sum u_r^{A_2} y_{rj}^{A_2}}{\sum v_i^A x_{ij}^A} \le 1, \quad \frac{\sum u_r^{B_1} y_{rj}^{B_1} + \sum u_r^{B_2} y_{rj}^{B_2}}{\sum v_i^B x_{ij}^B + \sum u_r^{A_1} y_{rj}^{A_1}} \le 1, \tag{9}$$

$$w_A \frac{\sum u_r^{A_1} y_{rj_0}^{A_1} + \sum u_r^{A_2} y_{rj_0}^{A_2}}{\sum v_i^A x_{ij_0}^A} + w_B \frac{\sum u_r^{B_1} y_{rj_0}^{B_1} + \sum u_r^{B_2} y_{rj_0}^{B_2}}{\sum v_i^B x_{ij_0}^B + \sum u_r^{A_1} y_{rj_0}^{A_1}} = \theta_{(AB)}^*, u_r^{A_1}, u_r^{A_2}, v_i^A, u_r^{B_1}, u_r^{B_2}, v_i^B, u_r^C, v_i^C \ge 0, j = 1, 2, ..., n.$$

In model (9), the retailer's efficiency is measured under the condition that the alliance's efficiency remains at $\theta_{(AB)}^*$. The optimal value $\theta_{(AB)C}^*$ is the retailer's efficiency when the alliance has first achieved its best performance. Each alliance member's efficiency also could be measured individually by using the optimal solution of model (9) as in (10).

$$\theta_{A/(AB)}^* = \frac{\sum u_r^{A_1*} y_{rj}^{A_1} + \sum u_r^{A_2*} y_{rj}^{A_2}}{\sum v_i^{A*} x_{ij}^A}, \theta_{B/(AB)}^* = \frac{\sum u_r^{B_1*} y_{rj}^{B_1} + \sum u_r^{B_2*} y_{rj}^{B_2}}{\sum v_i^{B*} x_{ij}^B + \sum u_r^{A_1*} y_{rj}^{A_1}}. \tag{10}$$

The efficiency of the whole supply chain could be defined as the arithmetic mean of the alliance's and retailer's efficiencies in (11).

$$E_{(AB)C} = w_1 \theta_{(AB)} + w_2 \theta_{(AB)C}, \text{ and } w_1 + w_2 = 1 \tag{11}$$

where w_1 and w_2 are similar decision maker specified weights. The alliance also could be built up between the manufacture and retailer.

THE COOPERATIVE MULTISTAGE SUPPLY CHAIN MODELS

This section considers the cooperative structure of supply chain, where all of the members are evaluated simultaneously. The cooperative model seeks to maximize the joint efficiency defined as the arithmetic mean of supplier's, manufacturer's and retailer's efficiencies in model (12).

$$\max\left[w_A \frac{\sum u_r^{A_1} y_{rj_0}^{A_1} + \sum u_r^{A_2} y_{rj_0}^{A_2}}{\sum v_i^A x_{ij_0}^A} + w_B \frac{\sum u_r^{B_1} y_{rj_0}^{B_1} + \sum u_r^{B_2} y_{rj_0}^{B_2}}{\sum v_i^B x_{ij_0}^B + \sum u_r^{A_1} y_{rj_0}^{A_1}} + w_C \frac{\sum u_r^C y_{rj_0}^C}{\sum v_i^C x_{ij_0}^C + \sum u_r^{B_1} y_{rj_0}^{B_1}} \right] = \theta_{(ABC)},$$

$$\text{s.t.} \quad \frac{\sum u_r^{A_1} y_{rj}^{A_1} + \sum u_r^{A_2} y_{rj}^{A_2}}{\sum v_i^A x_{ij}^A} \le 1, \frac{\sum u_r^{B_1} y_{rj}^{B_1} + \sum u_r^{B_2} y_{rj}^{B_2}}{\sum v_i^B x_{ij}^B + \sum u_r^{A_1} y_{rj}^{A_1}} \le 1, \frac{\sum u_r^C y_{rj}^C}{\sum v_i^C x_{ij}^C + \sum u_r^{B_1} y_{rj}^{B_1}} \le 1, \qquad (12)$$

$$w_A + w_B + w_C = 1,\ u_r^{A_1}, u_r^{A_2}, v_i^A, u_r^{B_1}, u_r^{B_2}, v_i^B, u_r^C, v_i^C \ge 0,\ j = 1, \ldots, n.$$

where w_A, w_B and w_C are also decision maker specified weights for supplier, manufacturer and retailer. Each weight also could be defined according to the portion of the total resources devoted to each member. The optimal value θ^*_{ABC} is the efficiency of the whole supply chain. Using the optimal solution, each supply chain member's efficiency could be measured individually in (13).

$$\theta^*_{A/(ABC)} = \frac{\sum u_r^{A_1*} y_{rj}^{A} + \sum u_r^{A_2*} y_{rj}^{A_2}}{\sum v_i^{A*} x_{ij}^A},\quad \theta^*_{B/(ABC)} = \frac{\sum u_r^{B_1*} y_{rj}^{B_1} + \sum u_r^{B_2*} y_{rj}^{B_2}}{\sum v_i^{B*} x_{ij}^B + \sum u_r^{A_1*} y_{rj}^{A_1}},\quad \theta^*_{C/(ABC)} = \frac{\sum u_r^{C*} y_{rj}^C}{\sum v_i^{C*} x_{ij}^C + \sum u_r^{B_1*} y_{rj}^{B_1}}. \qquad (13)$$

NUMERICAL EXAMPLE

The DEA models proposed above are demonstrated with a numerical example in this section. The supplier has three inputs, two normal outputs, and one intermediate output which are also the intermediate input of the manufacturer. The manufacturer has another two normal inputs, two normal outputs, and one intermediate output which are also the intermediate input of the retailer. The retailer has another two normal inputs , and three normal outputs. Table 1 reports the efficiencies obtained from these DEA models. E-ABC, E-CBA, E-(AB)C, E-(BC)A and E-(ABC) denote the efficiency scores under forward and backward non-cooperative structure, partial-cooperative structure 1 and 2, and cooperative structure. e-A, e-CBA, e-A/(AB), e-(BC)A, and e-A/(ABC) indicate the supplier efficiency scores in five different DEA models above. e-ABC, e-C, e-(AB)C, e-C/(BC), e-C/(ABC) denote the retailer efficiency scores in five different DEA models above. It can be seen that almost all supply chains are calculated as efficient with the overall efficiency in model (1), while its three members may not efficient under the non-cooperative, partial-cooperative or cooperative supply chain model. This is mainly because of the ignorance of the intermediate measures in model (1).

The whole supply chain efficiencies for each DMU under different models are compared. On average, the overall efficiency is highest, the cooperative efficiency is second highest, the efficiencies under non-cooperative and partial-cooperative structures are lower than the two above. And there are no evident differences between non-cooperative and partial-cooperative efficiencies. The average efficiencies of supply chain, supplier and retailer are compared and shown in Figure 2. On average, for both the supplier and retailer, the non-cooperative efficiency scores are higher than the cooperative efficiency scores and the partial-cooperative efficiency scores. And there are no evident differences between the forward and backward efficiency scores for both the supplier and retailer under the non-cooperative structure.

Table 1. Data of efficiencies for the multistage supply chain and its members

DMU		1	2	3	4	5	6	7	8	9	10
Overall efficiency		1.0000	1.0000	0.9752	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Forward non-cooperative structure	E-ABC	0.9553	0.8209	0.7506	0.9771	0.8994	0.9628	1.0000	0.9135	0.9398	1.0000
	e-A	1.0000	1.0000	0.4619	0.9313	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	e-ABC	0.8659	0.5839	0.9769	1.0000	0.8917	1.0000	1.0000	0.9211	1.0000	1.0000
Backward non-cooperative structure	E-CBA	0.9553	0.8209	0.7577	0.9771	0.8994	0.9628	1.0000	0.9135	1.0000	1.0000
	e-C	0.8659	0.5839	0.9769	1.0000	0.8917	1.0000	1.0000	0.9211	1.0000	1.0000
	e-CBA	1.0000	1.0000	0.4619	0.9313	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Partial-cooperative structure 1	E-(AB)C	0.9553	0.8190	0.7257	0.9634	0.9395	0.9256	1.0000	0.9711	1.0000	1.0000
	e-(AB)C	0.8659	0.4571	0.9769	1.0000	0.8917	1.0000	1.0000	0.9134	1.0000	1.0000
	e-A/(AB)	0.7015	1.0000	0.4595	0.9146	1.0000	0.8331	1.0000	1.0000	0.9986	1.0000
Partial-cooperative structure 2	E-(BC)A	0.9598	0.8437	0.7591	0.9771	0.8907	0.7928	1.0000	0.9258	1.0000	1.0000
	e-(BC)A	0.9202	1.0000	0.4619	0.9313	1.0000	0.5824	1.0000	1.0000	1.0000	1.0000
	e-C/(BC)	0.8659	0.5323	0.9769	1.0000	0.8899	1.0000	1.0000	0.9175	1.0000	1.0000
Cooperative structure	E-(ABC)	0.9795	1.0000	0.8006	1.0000	0.9551	0.8980	1.0000	1.0000	1.0000	1.0000
	e-A/(ABC)	0.7605	1.0000	0.4596	0.9138	1.0000	0.8331	1.0000	1.0000	1.0000	1.0000
	e-C/(ABC)	0.8659	0.5215	0.8006	1.0000	0.8917	1.0000	1.0000	0.7497	1.0000	1.0000

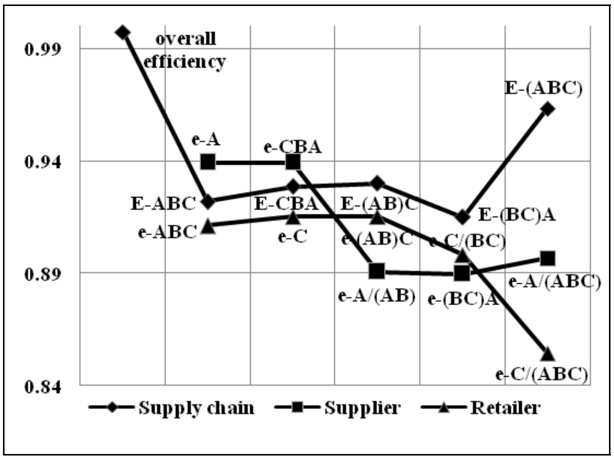


Figure 2. Average efficiency scores

CONCLUSION

In this paper, we developed several multistage supply chain DEA models for measuring the efficiencies of the whole supply chain and it members under different supply structures. We consider a representative multistage supply chain which has three members: supplier, manufacturer and retailer. The non-cooperative structure model is first proposed to evaluate the efficiency of each supply chain member according to a specified sequence. The partial-cooperative structure model is second proposed to evaluate the efficiency of the alliance, which is constructed between two of the supply chain members, and the other member which is not in the alliance. The cooperative structure model is third proposed to evaluate the efficiencies of all of the supply chain members simultaneously. The intermediate measures between each supply chain member are considered which link the members and the whole supply chain. Two different kinds of inputs/outputs are also considered in the models in order to make the multistage supply chain models becoming more general. The models proposed in this paper could be seen as efficient tools for the performance evaluation of multistage supply chain. We demonstrate these models with a numerical example, which illustrate that the multistage models could provide more detailed information about the efficiency or inefficiency decomposition to each supply member. We have to point out that the models proposed in this paper are based on constant returns to scale (CRS) DEA. The variable returns to scale (VRS) DEA for multistage supply chain performance evaluation is a subject for future research.

REFERENCE

[1] Charnes, A. Cooper, W. W. Rhodes, E. “Measuring the efficiency of decision making units.” European Journal of Operational Research, 1978, 2(6), 429-444.

[2] Banker, R. D. Charnes, A. Cooper, W. W. “Some models for the estimation of technical and scale inefficiencies in data envelopment analysis.” Management Science, 1984, 30(9), 1078-1092.

[3] Cooper, W. W. Seiford, L. M. Zhu, J. Handbook on data envelopment analysis. Boston, MA: Kluwer Academic Publishers, 2004.

[4] Seiford, L. M. Zhu, J. “Profitability and marketability of the top 55 US commercial banks.” Management Science, 199, 45(9), 1270-1288.

[5] Zhu, J. “Multi-factor performance measure model with an application to Fortune 500 companies.” European Journal of Operational Research, 2000, 123(1), 105-124.

[6] Fare, R. Grosskopf, S. “Network DEA.” Socio-Economic Planning Sciences, 2000, 34(1), 35-49.

[7] Sexton, T. R. Lewis, H. F. “Two-stage DEA: an application to major league baseball.” Journal of Productivity Analysis, 2003, 19(2-3), 227-249.

[8] Chen, Y. Zhu, J. “Measuring information technology’s indirect impact on firm performance.” Information Technology and Management Journal, 2004, 5(1-2), 9-22.

[9] Kao, C. Hwang, S. N. “Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan.” European Journal of Operational Research, 2008, 185(1), 418-429.

[10] Chen, Y. Cook, W. D. Li, N. Zhu, J. “Additive efficiency decomposition in two-stage DEA.” European Journal of Operational Research, 2009, 196(3), 1170-1176.

[11] Liang, Yang, L. F. Cook, W. D. Zhu, J. “DEA models for supply chain efficiency evaluation.” Annals of Operations Research, 2006, 145(1), 35-49.

EXPLORING THE ENABLERS OF SUPPLIER RESPONSIVENESS IN B2B SUPPLY CHAINS

Tobin E. Porterfield, Towson University,
8000 York Road, Towson, MD 21252, tporterfield@towson.edu, 410-704-3265

Chaodong Han, Towson University,
8000 York Road, Towson, MD 21252, chan@towson.edu, 410-704-4658

Dmitriy Nesterkin, Towson University,
8000 York Road, Towson, MD 21252, dnesterkin@towson.edu, 410-704-3031

ABSTRACT

This study represents a significant contribution to the under-researched area of supplier responsiveness in providing information to industrial customers. The exchange of information supports decision making throughout the supply chain and is an often overlooked component of supplier responsiveness. This study brings together research from multiple disciplines to develop hypotheses and a research model to identify conditions and capabilities that enable supplier responsiveness to requests for information. The proposed model and methodology are provided as part of this ongoing research project.

Keywords: Supply chain, information exchange, supplier responsiveness

INTRODUCTION

There is considerable research on the importance of the interactions of firms within the context of supply chains. These networks of buyer-supplier relationships are shown to either enable or thwart a firm's ability to meet the needs of its customers. Research supports that integration with suppliers enhances firm performance [1] and that the level of integration varies depending on the specific nature of the relationship [2, 3]. Additionally, firms that are more integrated internally and externally show increased capabilities to provide a broad range of solutions to their customers [4].

Research on supplier development programs (SDP) has addressed communication strategies [5-7], knowledge transfer [8], and supplier dependence [9]. While these studies improve the understanding of buyer efforts in SDP, they have not adequately addressed the capabilities that enable a supplier to meet the buying firm's expectations. Underlying any buyer-supplier business transaction is the exchange of information. The supplier must respond with the information which will allow the customer to make operational decisions. This study focuses on supplier responses to requests for information from their industrial customers. The information provided by the supplier is often a critical input to the customer firm's decision making process and affects their ability to serve their customers.

While existing research has recognized the critical role of supplier responsiveness in supply chain performance, minimal research has addressed the sources of supplier responsiveness. This paper proposes a framework and methodology to empirically answer the following research question:

What are conditions and capabilities that enable supplier responsiveness?

LITERATURE REVIEW AND HYPOTHESES

This study draws on the supplier development, operations management, marketing, and logistics / supply chain management literature to build a comprehensive model of supplier responsiveness. This section begins by defining supplier responsiveness across the disciplines and is followed by identification of potential sources of supplier responsiveness.

Defining Supplier Responsiveness

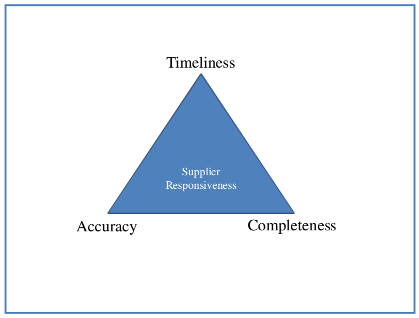
The term *supplier responsiveness* has been used to describe many types of reactions by firms supplying goods and services. As the area of buyer-supplier interactions has matured, many types of supplier reactions to customer requests have moved out from under the general performance area of supplier responsiveness and now stand on their own as distinct supplier capabilities (e.g. agility, flexibility, adaptability).

Logistics research offers a two-part definition of responsiveness in a supply chain context. This definition makes a distinction based on the direction of flow. The first part focuses on the timeliness of the flow of information and goods while the second part addresses the transparency of product demand information [10]. The product demand information flows upstream towards the manufacturer in order to support forecasting and production planning. Additional information and the physical product flow downstream toward the end customer. Since this study is deliberately restricted to supplier responsiveness to information requests from their industrial customers, the information flow is of particular interest. The flow of the physical product is not included in this study and is more appropriately addressed as supply chain flexibility or how supply chains adapt to disruptions and changes in demand [11]. The ability of the supply chain to adjust the physical flow of goods is also addressed in the literature as flexibility [11], agility [12], agile manufacturing [13], and manufacturing flexibility [14]. This study maintains a focus on supplier responsiveness as part of a time-based strategy related to the exchange of information within buyer-supplier relationships.

Additional aspects of responsiveness are found when it is recognized that information has both time and content dimensions. Supplier responsiveness has been defined to include both the timeliness and accuracy of a supplier's response to a request from their customer [17]. Through a study of technology-enabled information exchange, the content dimension is expanded to recognize the attributes of accuracy and completeness [18]. Also, timeliness of service is considered one of the key determinants of responsiveness from a customer perspective [19].

The inclusion of both time and content dimensions in the definition of supplier responsiveness more accurately captures the role of supplier responses in the decision making process. If only the time dimension is included when evaluating supplier responsiveness then quick responses that are inaccurate or incomplete would be considered valuable. Clearly, it is in the best interest of the buying firm to use supplier responses that are timely, accurate, and complete. As shown in Figure 1, supplier information responsiveness is inclusive of three dimensions: *the ability of a firm to provide timely, accurate, and complete information to their customers.*

FIGURE 1: DIMENSIONS OF SUPPLIER RESPONSIVENESS



Hypothesis Development

Three primary research areas were examined in order to identify potential enablers of supplier responsiveness and support the development of an appropriate analytical framework. Research in the areas of supplier development, marketing, and logistics / supply chain management were examined to identify the underlying factors that enable supplier responsiveness. Table 1 summarizes the proposed hypotheses.

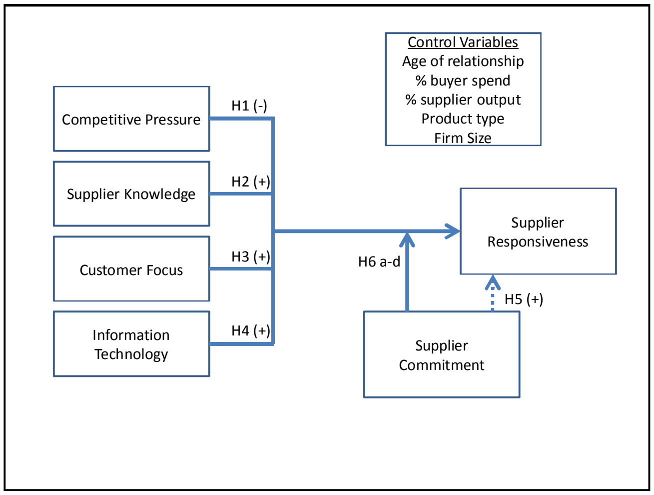
TABLE 1: HYPOTHESES

Area		Hypothesis
Competitive Pressure	H1	*Competitive pressure is negatively associated with supplier responsiveness*
Supplier Knowledge	H2	*Supplier knowledge is positively associated with supplier responsiveness*
Customer Focus	H3	*Customer focus is positively associated with supplier responsiveness*
Use of Information Technology	H4	*The use of information technology is positively associated with supplier responsiveness*
Supplier Commitment	H5	*Supplier commitment is positively associated with supplier responsiveness*
	H6a	*Supplier commitment moderates the relationship between competitive pressure and supplier responsiveness*
	H6b	*Supplier commitment moderates the relationship between supplier knowledge and supplier responsiveness*
	H6c	*Supplier commitment moderates the relationship between customer focus and supplier responsiveness*
	H6d	*Supplier commitment moderates the relationship between the use of information technology and supplier responsiveness*

RESEARCH MODEL AND METHODOLOGY

The research model provided in Figure 2 shows the hypothesized relationships. On the left are the internal and external factors that affect supplier responsiveness. The external component is the *competitive pressure* created by the buying firm's sourcing strategy. The remaining three factors on the left represent capabilities internal to the supplier. Additionally, *supplier commitment* is hypothesized to have both a direct effect on supplier responsiveness and a moderating effect on the other four factors.

FIGURE 2: RESEARCH MODEL



Methodology

This study proposes to collect data through a two-stage survey methodology. An initial survey of industrial customers would be distributed requesting that they identify and rate five of their first-tier suppliers. A second survey would be sent to the identified suppliers to collect information on their firm characteristics and capabilities [20, 21]. The matched perceived supplier responsiveness measures and supplier characteristics will provide a robust analysis of the enablers of supplier responsiveness.

CONCLUSION AND NEXT STEPS

From a research perspective, this study represents the first known effort to bring together multiple sources of supplier information responsiveness for testing in a single statistical model. For practitioners, the results should assist buying firms in identify suppliers that have capabilities to support the required level of supplier responsiveness. Additionally, the results will guide suppliers in developing information responsiveness capabilities.

Acknowledgement: This research was funded through a Summer Research Fellowship by the Faculty Development and Research Committee of Towson University.

REFERENCES

1. Frohlich, M.T. and R. Westbrook, "Arcs of integration: an international study of supply chain strategies". *Journal of Operations Management*, 2001. 19: p. 185-200.
2. McCutcheon, D.M. and F.I. Stuart, "Issues in the choice of supplier alliance partners". *Journal of Operations Management*, 2000. 18: p. 279-301.
3. Lambert, D.M., M.A. Emmelhainz, and J.T. Gardner, "So you want to partner?" *Marketing Management*, 1996. 5(2): p. 24-41.
4. Daugherty, P.J., R.E. Sabath, and D.S. Rogers, "Competitive advantage through customer responsiveness". *Logistics and Transportation Review*, 1992. 28(3): p. 257-271.
5. Prahinski, C. and W.C. Benton, "Supplier evaluations: Communication strategies to improve supplier performance". *Journal of Operations Management*, 2004. 22(1): p. 39-62.
6. Prahinski, C. and Y. Fan, "Supplier evaluations: The role of communications quality". *The Journal of Supply Chain Management*, 2007. 43(3): p. 16-28.
7. Joshi, A.W., "Continuous supplier performance improvement: Effects of collaborative communication and control". *Journal of Marketing*, 2009. 73: p. 133-150.
8. Modi, S.B. and V.A. Mabert, "Supplier development: Improving supplier performance through knowledge transfer". *Journal of Operations Management*, 2007. 25: p. 42-64.
9. Carr, A.S., H. Kaynak, J.L. Hartley, and A. Ross, "Supplier dependence: Impact on supplier's participation and performance". *International Journal of Operations & Production Management*, 2008. 28(9): p. 899-916.
10. Catalan, M. and H. Kotzab, "Assessing the responsiveness in the Danish mobile phone supply chain". *International Journal of Physical Distribution and Logistics Management*, 2003. 33(8): p. 668-685.
11. Stevenson, M. and M. Spring, "Flexibility from a supply chain perspective: Definition and review". *International Journal of Operations & Production Management*, 2007. 27(7): p. 685-713.
12. Christopher, M., "The agile supply chain: Competing in volatile markets". *Industrial Marketing Management*, 2000. 29: p. 37-44.
13. Perry, M. and A.S. Sohal, "Effective quick response practices in a supply chain partnership: An Australian case study". *International Journal of Operations & Production Management*, 2000. 21(5/6): p. 840-54.
14. Pagell, M. and D. Krause, "A multiple-method study of environmental uncertainty and manufacturing flexibility". *Journal of Operations Management*, 1999. 17: p. 307-15.
15. Wang, E.T.G., J.C.F. Tai, and H.-L. Wei, "A virtual integration theory of improved supply chain performance". *Journal of Management Information Systems*, 2006. 23(2): p. 41-64.
16. Carr, A.S. and L.R. Smeltzer, "An empirical study of the relationships among purchasing skills and strategic purchasing, financial performance, and supplier responsiveness". *Journal of Supply Chain Management*, 2000. 36(3): p. 40-54.
17. Choi, T.Y. and D. Krause, "The supply base and its complexity: Implications for transaction costs, risks, responsiveness, and innovation". *Journal of Operations Management*, 2006. 24: p. 637-652.

18. Sellitto, C., S. Burgess, and P. Hawking, "Information quality attributes associated with RFID-derived benefits in the retail supply chain". *International Journal of Retail and Distribution Management*, 2007. 35(1): p. 69-87.
19. Parasuraman, A., V.A. Zeithaml, and L.L. Berry, "A conceptual model of service quality and its implications for future research". *Journal of Marketing*, 1985. 49 (Fall): p. 41-50.
20. Hult, G.T.M., D.J. Ketchen Jr., and E.L. Nichols Jr., "Organizational learning as a strategic resource in supply management". *Journal of Operations Management*, 2003. 21: p. 541-556.
21. Vickery, S.K., C. Droge, T.P. Stank, T.J. Goldsby, and R.E. Markland, "The performance implications of media richness in a business-to-business service environment: direct versus indirect effects". *Management Science*, 2004. 50(8): p. 1106-1119.

A STUDY ON THE COORDINATION OF THE TWO-ECHELON SUPPLY CHAIN USING CREDIT AND QUANTITY DISCOUNT OPTIONS

Ruo Du, Drexel University, (215) 313-3423, rd344@drexel.edu
Seung-Lae Kim, Drexel University, (215) 895-2181, kimsl@drexel.edu

ABSTRACT

In this paper we study the coordination of the two-echelon supply chain using credit and quantity discount options. We determine the retail price, order quantity for the buyer, and the setup frequency, credit period for the supplier in four different scenarios (with or without coordination and with or without credit option). And we suggest a policy that incorporates both quantity discount and credit option, which will benefit both parties and maximize the total supply chain profit.

INTRODUCTION

In the last twenty years, supply chain coordination has been receiving enormous attention as it can achieve a global optimality for the supply chain and benefit all related parties. Some mechanisms such as quantity discounts, credit option, buy back, etc have been used for coordination purpose. In this paper, we compare the two coordination mechanism-quantity discount and credit option by setting and analyzing four scenarios (with or without coordination and with or without credit option). We build the models with price-sensitive demand, production rate that is greater than demand rate, for a single production multiple delivery system. We determine the retail price, order size, production size and credit period through mechanism to achieve an optimal solution for both parties individually or for the integrated system simultaneously. Then we explain where the additional profit originates from, and give advice on making decisions of when to use credit as coordination mechanism. We also suggest how to set up an appropriate policy to fully utilize the coordination, and how to divide the extra profit, which are incurred by both quantity discount and credit option.

NOTATIONS AND ASSUMPTIONS

Supplier:

v: Unit price charged by supplier to the buyer
m: Manufacturing cost of the supplier
n: Number of batches per production run
R: Annual Production rate
ρ: Capacity utilization where $\rho = D/R \text{ and } 0 < \rho \leq 1$
A_1: Setup cost per setup
h_1^c: Capital cost of the supplier
h_1^s: Other inventory cost of the supplier
h_1: Holding cost of the supplier where $h_1 = h_1^c + h_1^s$
Π_1: Annual profit of the supplier

Buyer:

p: Retail price charged by the buyer
D: Annual demand rate, function of the retail price
Q: Order quantity of the buyer
A_2: Ordering cost per order
h_2^c: Capital cost of the buyer
h_2^s: Other inventory cost of the buyer
h_2: Holding cost of the buyer where $h_2 = h_2^c + h_2^s$
Π_2: Annual profit of the buyer

Chain:
t: Credit time
Π: Profit of the chain
d: Quantity discount factor
I: Interest charged for the credit period
h_d: Price charged per item for the credit period, where $h_d = I * v$

Assumptions:
1. Only one supplier and one buyer are considered in this study. They share complete information. And both parties strictly follow the contracts.
2. Production rate is greater than demand rate.
3. Shortages are not allowed.
4. The other inventory costs excluding capital costs are the same for both parties.
5. The capital holding cost $h_i^c = r_i \cdot v$ ($i = 1,2$ where $i = 1$(supplier) and $i = 2$(buyer)), in which r_i is the capital cost per dollar per unit time. h_i^c is different for different items although r_i is the same for all items.
6. Annual demand is price sensitive, $D(p) = kp^{-\beta}$ where β is the elasticity coefficient. k is a constant coefficient of the demand function.
7. The manufacturer produces a n integer multiple of the quantity the buyer ordered.
8. Lead time is zero for replenishments.

NON-COORDINATION WITHOUT CREDIT

In this section we consider the situation where no coordination, neither credit nor quantity discount is present. The buyer optimizes its profit by deciding order size and retail price. And the vender decides the batch multiplier n that maximizes its profit. The buyer decides the retail price and order quantity (p, Q) to maximize its profit. The annual profit function is composed of sales revenue, production cost, setup cost and holding costs, including storage and capital costs. We can find the optimal price that maximizes the buyer's profit. The vendor here only considers the number of batches produced each setup and maximizes the profit. We can get the optimal number of batches n that maximizes the vendor's profit.

NON-COORDINATION WITH CREDIT

This is the scenario where the buyer optimizes its profit by determining order size Q and retail price p and the vendor optimizes its profit by choosing both credit time t and the number of batches n, assuming he knows how the buyer will decide price and order quantity based on credit provided to achieve optimality.

COORDINATION WITH AND WITHOUT CREDIT

Under the assumption that the supplier and buyer share complete information and both strictly follow the credit contract, we compute here the total supply chain profit by determining the retail price and lot size for the buyer, the multiple of order quantity per setup for the supplier. Since the procedure and algorithms for the scenario of coordination without credit will be the same with the scenario with credit, we here only explore one model. The results for the scenario without credit will be similar to the results below only without the credit value factor $(h_2^c - h_1^c) \cdot Dt$.

COORDINATION MECHANISM SELECTION AND PROFIT SHARING

In this section we discuss how to share the additional value gained through coordination so that both parties have a reasonable increase in profit, compared to the non-coordination scenario without credit. When the capital cost for the supplier is greater than that of the buyer $h_2^c < h_1^c$. It is more economical for the chain to adopt a quantity discount option than to use the credit option [20]. If the vendor wants to stimulate the end demand when the buyer has not enough cash in hand, the vendor may want to use a credit option. The supplier can allow the buyer to have some permissible time until payment at an interest rate that is between their capital holding costs $h_2^c < h_d = I * v < h_1^c$. In this case, the chain can benefit not only from the quantity discount coordination but also from the difference in capital holding costs.

In coordination without credit, we use quantity discount to divide the profit. In coordination with credit option when $h_1^c < h_2^c$, the longer the credit period, the more profits the supply chain gains. Thus the vendor will encourage the buyer to use the credit option. In sum, the supplier uses first the quantity discount option and then the credit option. To keep the credit option without the supplier losing money, he should charge the buyer interest on the amount owned at a rate that is between their capital holding costs.

NUMERICAL EXAMPLE AND SENSITIVITY ANALYSIS

It can be seen that when credit is adopted, the supply chain profit increases a little, because it stimulates the buyer to order more each time. When two parties act in a coordinated fashion, total supply chain profit increases significantly. Order quantity increases and retail price decreases significantly (from about 30 to about 15.6). Also the supplier can produce more times the order size of the buyer (6 to 7) in scenario of $\rho = 0.7$. And total chain profit increased by about 12%. When we use credit option as coordination mechanism, this increase grows even larger.

CONCLUSION

In this paper, we compare two different coordination models under various scenarios to explore the origin and incentive for extra profit. Based on that, we can develop appropriate polices and contracts to achieve as much profits for the chain as possible. Reasonable and easy carrying profit sharing methods are suggested where both quantity discount and credit option are used. In this case, the optimal chain profit is gained under suitable and comfortable credit length for both parties. Profits can also be easily shared under negotiation with flexibility. The suggested policy can fully utilize the drive for additional profit from the difference in capital cost in credit, and it is also very easy to apply. One other advantage of this method is that the credit length can be very flexible compared to that in the scenario of using credit option alone. The price, quantity size and number of batches produced do not need to be recomputed for new optimization when credit period is changed. The profit is shared nicely in the form of interest rate charged by the supplier, which makes the policy easy to be carried out in flexible manner in practice.

REFERENCES

[1] Abad and Jaggi, 2003 P.L. Abad and C.K. Jaggi, A joint approach for setting unit price and the length of the credit period for a seller when end demand is price sensitive, *International Journal of Production Economics* **83** (2003) (2), pp. 115–122.

[2] Banerjee, 1986 A. Banerjee, A joint economic-lot-size model for purchaser and vendor, *Decision Sciences* **17** (1986), pp. 292–311.

[3] Chang et al., 2006 H.C. Chang, L.Y. Ouyang, K.S. Wu and C.H. Ho, Integrated vendor–buyer cooperative inventory models with controllable lead time and ordering cost reduction, *European Journal of Operational Research* **170** (2006), pp. 481–495.

[4] Chang et el.(2009) LY Ouyang, HC Chang, CH Ho, CH Su, The optimal pricing and ordering policy for an integrated inventory model when trade credit linked to order quantity, APPLIED MATHEMATICAL MODELLING **33** (7) JUL 2009, pp. 2978-2991

[5] Chen and Chen, 2005 T.H. Chen and J.M. Chen, Optimizing supply chain collaboration based on joint replenishment and channel coordination, *Transportation Research Part E: Logistics and Transportation Review* **41** (2005), pp. 261–285.

[6] Chen and Kang, 2007 LH Chen and FS Kang, Integrated vendor-buyer cooperative inventory models with variant permissible delay in payments, *EUROPEAN JOURNAL OF OPERATIONAL RESEARCH*, **183** (2), pp 658-673

[7] Chung, 1998 K.J. Chung, A theorem on determination of economic order quantity under conditions of permissible in payments, *Computers and Operations Research* **25** (1998) (1), pp. 49–52.

[8] Goyal, 1976 S.K. Goyal, An integrated inventory model for a single supplier-single customer problem, *International Journal of Production Research* **15** (1) (1976), pp. 107–111.

[9] Goyal, 1985 S.K. Goyal, Economic order quantity under conditions of permissible delay in payments, *Journal of the Operational Research Society* **36** (1985), pp. 335–338.

[10] Goyal, 1988 S.K. Goyal, A joint economic-lot-size model for purchaser and vendor: A comment, *Decision Sciences* **19** (1988), pp. 236–241.

[11] Harris, 1915 Harris, F. W. (1915). *Operations and Cost-Factory Management Series* (A.W. Shaw Co., Chicago), chapter 4.

[12] Jaber and Osman, 2006 M.Y. Jaber and I.H. Osman, Coordinating a two-level supply chain with delay in payments and profit sharing., *Computers & Industrial Engineering* **50** (4) (2006), pp. 385–400.

[13] Khouja and Mehrez, 1994 M. Khouja and A. Mehrez, Economic production lot size model with variable production rate and imperfect quality, *Journal of the Operational Research Society* **45** (1994), pp. 1405–1417.
[14] Kim et al., 1995 S.J. Kim, H. Hwang and W.S. Shinn, An optimal credit policy to increase wholesaler's profit with price dependent demand function, *Production Planning and Control* **6** (1995) (1), pp. 45–50.
[15] Lee and Rosenblatt, 1986 H.L. Lee and M.J. Rosenblatt, A generalized quantity discount-pricing model to increase supplier's profits, *Management Science* **32** (1986) (9), pp. 1177–1185.
[16] Mehta, 1968 D. Mehta, The formulation of credit policy models, *Management Science* **15** (1968) (2), pp. B30–B50.
[17] Munson and Rosenblatt, 2001 L.C. Munson and J.M. Rosenblatt, Coordinating a three level supply chain with quantity discounts, *IIE Transactions* **33** (2001) (4), pp. 371–384.
[18] Ouyang, Ho, Su (2008) Optimal strategy for an integrated system with variable production rate when the freight rate and trade credit are both linked to the order quantity, INTERNATIONAL JOURNAL OF PRODUCTION ECONOMICS **115** (1) SEP 2008, pp. 151-162
[19] Pan and Yang, 2002 J.C.-H. Pan and J.S. Yang, A study of an integrated inventory with controllable lead time, *International Journal of Production Research* **40** (2002), pp. 1263–1273.
[20] Sarmah et al., 2007 S.P. Sarmah, D. Acharya and S.K. Goyal, Coordination and profit sharing between a manufacturer and a buyer with target profit under credit option, *European Journal of Operational Research* **182** (3) (2007), pp. 1469–1478.
[21] Sheen and Tsao, 2007 G.J. Sheen and Y.C. Tsao, Channel coordination, trade credit and quantity discounts for freight cost, *Transportation Research Part E: Logistics and Transportation Review* **43** (2007), pp. 112–128.
[22] Shinn and Hwang, 2003 W.S. Shinn and H. Hwang, Optimal pricing and ordering policies for retailers under order size dependent delay in payments, *Computers and Operations Research* **30** (2003) (1), pp. 35–50.
[23] Yang and Wee, 2006 P.C. Yang and H.M. Wee, A collaborative inventory system with permissible delay in payment for deteriorating items, *Mathematical and Computer Modeling* **43** (3–4) (2006), pp. 209–221.

DEMANDS ALONG THE SUPPLY CHAIN

ABSTRACT

This paper describes how the monthly demands vary at the locations along the supply chain, coming from the customers to a dealer onto a distribution center and finally to a supplier. The monthly demands at the dealer are assumed to follow the Poisson probability distribution. The mean, standard deviation and coefficient of variation are measured for each of the locations.

INTRODUCTION

Consider a dealer that carries inventory on parts to meet the oncoming customer demands. This paper assumes the monthly demands from the customers are horizontal (no trend or seasonal pattern); and also, the monthly demands at a dealer are shaped like a Poisson probability distribution. The dealer replenishes the stock on the part from a distribution center (DC), and the DC is replenished from a supplier. The replenish quantities Q at a location depends mostly on the forecast of demands and the cost per unit. The replenish order quantity can be stated in month-in-buy (mib) terms where $Q = mib \times \mu$ and μ is the forecast of the average monthly demands at the location. When mib is 1.0, Q is one month supply; when mib = 2.0, Q is two months supply, and so forth. The aggregate flow of replenishments from the dealers are the demands at the DC's; and the flow of demands from the DC's are the demands at the supplier. This paper shows how the demands along the supply chain, at the dealer, DC and the supplier are related.

ALL CUSTOMERS MONTHLY DEMANDS

Assume an item where μ_0 is the average monthly demand from all customers. This paper gives examples where μ_0 = 10, 100 and 1000.

AN AVERAGE DEALER MONTHLY DEMANDS

The demand for an average dealer is denoted as d_1 and when N_1 is the number of dealers, the mean monthly demand for an average dealer is μ_1, and $\mu_1 = \mu_0/N_1$.

This paper assumes the monthly demands to a dealer from its customers is Poisson distributed with an average of μ_1. Since the monthly demands are Poisson, the associated standard deviation is $\sigma_1 = \sqrt{\mu_1}$. Recall, the coefficient of variation (cov) of a random variable is defined as the ratio of the standard deviation over the average. So in this situation, the coefficient of variation becomes $cov_1 = \sigma_1/\mu_1$. Table 1 lists values μ_1 ,σ_1 and cov_1 as the average monthly demands range from 1 to 1000. Note how cov_1 is largest (1.00) when $\mu_1 = 1$ and becomes increasingly smaller as μ_1 rises. Further, when the demands are all positive and normally distributed, the cov attains a value is 0.33 or less. This is because the mean is at least three standard deviations larger than zero. On the other extreme, when the cov is one, the monthly demands are distributed as an exponential distribution, since for this distribution, the standard deviation is the same as the mean. So in essence, as the average monthly demands (from customers) increase, the monthly demands tends like a *normal* distribution; and as the average monthly demands go down, the monthly demands are of the lumpy type and shaped more like an *exponential* distribution.

Table 1. Monthly demand statistics for an average dealer

μ_1	σ_1	cov_1
1	1.00	1.00
5	2.24	0.45
10	3.16	0.32
50	7.07	0.14
100	10.00	0.10
500	22.36	0.04
1000	31.62	0.03

AN AVERAGE DC MONTHLY DEMANDS

The monthly demands for an average distribution center (DC) is here denoted as d_2. The associated mean monthly demand is μ_2 and corresponding standard deviation is σ_2. Also, the coefficient of variation is $cov_2 = \sigma_2/\mu_2$. The compute the measures (μ_2 , σ_2, cov_2) for an average DC, the analysis below is presented.

ANALYSIS

When the dealer needs replenish stock, they buy from its assigned distribution center (DC). The buy quantities are in lot sizes that are economical for them. Suppose the lot size is q and this quantity is sufficient for the forecast of demands over the next mib months. In this analysis, mib = months-in-buy and represents the buy amount of replenish stock in monthly requirements. Of interest here is to measure the mean and variance for the average monthly replenish quantities from the dealer to the DC. For notation, E(q) = expected replenish quantity per month and V(q) is the associated variance.

Consider the general situation when the months-in-buy, mib, is an integer of m = 1,2, 3,... Suppose the dealer demands for the most recent m months are: d(1), ..., d(m). Also assume the stock at the dealer is adequate at the first month (t = 0), and the dealer needs replenish stock at month t = m, for a quantity size that covers the next m month requirements. So the replenish quantity from the dealer to the DC will be q(t) = 0 for months t = 1 to m-1 and will be approximately q(m) = (d(1)+...+d(m)) at the end of month m. Since mib = m, this pattern will repeat as the months move along. Of interest now is to determine the expected value and variance of the monthly replenish quantities, q(t). Note (m-1) of the quantities are zero, and one quantity is an m-month supply. The expected value of q is:

$$
\begin{aligned}
E(q) &= ((m-1)/m) \times 0 + 1/m \times E(d(1) + ... + d(m)) \\
&= 1/m \times m\, E(d) \\
&= E(d).
\end{aligned}
$$

In a similar way, the expected value of q^2 becomes

$$
\begin{aligned}
E(q^2) &= 1/m\, E[(d(1)+...+d(m))^2] \\
&= 1/m(mE(d^2) + m(m-1)E(d)^2) \\
&= E(d^2) + (m-1)E(d)^2 \\
&= V(d)+E(d)^2 + (m-1)E(d)^2 \\
&= V(d) + mE(d)^2.
\end{aligned}
$$

Finally, the variance of q becomes

$$
\begin{aligned}
V(q) &= E(q^2) - E(q)^2 \\
&= V(d)+(m-1)E(d)^2.
\end{aligned}
$$

Returning to the notation for the average DC, the following substitutions are made: $\mu_2 = E(q)$, $\sigma_2^2 = V(q)$, $\mu_1 = E(d)$ and $\sigma_1^2 = V(d)$. The month-in-buy from the dealer to the DC is denoted as M_1. Further, N_2 is the number of dealers and N_3 the number of DC's. The statistics for the mean monthly demands d_2 at an average DC are as below:

$$\mu_2 = (N_2/N_3)\mu_1$$
$$\sigma_2^2 = (N_2/N_3)(\sigma_1^2 + [M_1 - 1]\mu_1^2)$$
$$cov_2 = c_2 = \sigma_2/\mu_2.$$

Table 2 shows how cov_1 and M_1 = mib are related to cov_2 as cov_1 ranges from (0.1 to 1.0) and M_1 from (1 to 6). Recall, cov_1 is a measure of the variation in demands from the customers to an average dealer, and cov_2 is the counterpart measure of the variation in demands from the dealer to the an average DC . Note how cov_2 increases significantly as the M_1 increases beyond one month. Also note when M_1 is two or larger, cov_2 increases gradually as cov_1 goes from 0.1 to 1.0.

Table 2. Values of cov_2 as related to cov_1 and M_1

M_1 \ cov_1	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
1	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
2	1.00	1.02	1.04	1.08	1.01	1.17	1.22	1.28	1.35	1.41
3	1.42	1.43	1.45	1.47	1.50	1.54	1.58	1.62	1.68	1.73
4	1.73	1.74	1.76	1.78	1.80	1.83	1.87	1.91	1.95	2.00
5	2.00	2.01	2.02	2.04	2.06	2.09	2.12	2.15	2.19	2.24
6	2.24	2.24	2.26	2.27	2.29	2.32	2.34	2.37	2.41	2.45

AVERAGE MONTHLY DEMANDS FOR ALL DC's

In this section, assume N_3 is the number of DC's and d_3 is the average demand for all DC's. Further μ_3 is the average monthly demand, σ_3 the standard deviation and c_3 is the coefficient of variation of d_3. So now, $\mu_3 = N_3 \times \mu_2$, $\sigma_3^2 = N_3 \times \sigma_2^2$ and $cov_3 = c_3 = \sigma_3 / \mu_3$.

AVERAGE MONTHLY DEMANDS FOR THE SUPPLIER

The monthly demands for the supplier is denoted as d_4. In the same way, μ_4 is the associated average monthly demand, σ_4 the standard deviation and c_4 is the coefficient of variation of d_4. In the calculations, M_3 is the month-in-buy from the DC to the supplier.

So now,
$\mu_4 = \mu_3$
$\sigma_4^2 = \sigma_3^2 + (M_3 - 1)\mu_3^2$
$c_4 = \sigma_4/\mu_4$.

SUMMARY STATISTICS ACROSS THE SUPPLY CHAIN

In summary, the demand statistics across the supply chain are given in Tables 3.1 and 3.2. To review, the notation used is listed below:

N_2 = number of dealers
N_3= number of DC's
M_1 = dealer month-in-buy (to DC)
M_3 = DC month-in-buy (to supplier)
d_0 = monthly demands all customers
d_1 = monthly customer demands per dealer
μ_1, σ_1 are the mean and standard deviation of d_1
c_1 = cov of d_1
d_2 = monthly demands per DC
μ_2, σ_2 are the mean and standard deviation of d_2
c_2 = cov of d_2
d_3 = monthly demands all DCs

μ_3, σ_3 are the mean and standard deviation of d_3
c_3 = cov of d_3
d_4 = monthly demands to supplier
μ_4, σ_4 are the mean and standard deviation of d_4
c_4 = cov of d_4

Table 3.1, Average monthly demands (μ) and cov (c) for all customers, per dealer, per DC, all DC's and supplier when number of dealers is N_2 = 100 and number of DC's is N_3 = 1.

		All	per dealer		per DC		all DCs		supplier	
M_1	M_3	μ_0	μ_1	c_1	μ_2	c_2	μ_3	c_3	μ_4	c_4
1	1	10	0.1	3.16	10.0	0.32	10.0	0.32	10.0	0.32
1	1	100	1.0	1.00	100.0	0.10	100.0	0.10	100.0	0.10
1	1	1000	10.0	0.32	1000.0	0.03	1000.0	0.03	1000.0	0.03
1	2	10	0.1	3.16	10.0	0.32	10.0	0.32	10.0	1.05
1	2	100	1.0	1.00	100.0	0.10	100.0	0.10	100.0	1.00
1	2	1000	10.0	0.32	1000.0	0.03	1000.0	0.03	1000.0	1.00
2	1	10	0.1	3.16	10.0	0.33	10.0	0.33	10.0	0.33
2	1	100	1.0	1.00	100.0	0.14	100.0	0.14	100.0	0.14
2	1	1000	10.0	0.32	1000.0	0.10	1000.0	0.10	1000.0	0.10
2	2	10	0.1	3.16	10.0	0.33	10.0	0.33	10.0	1.05
2	2	100	1.0	1.00	100.0	0.14	100.0	0.14	100.0	1.01
2	2	1000	10.0	0.32	1000.0	0.10	1000.0	0.10	1000.0	1.01

Table 3.2, Average monthly demands (μ) and cov (c) for all customers, per dealer, per DC, all DC's and supplier when number of dealers is N_2 = 1000 and number of DC's is N_3 = 5.

		All	per dealer		per DC		all DCs		supplier	
M_1	M_3	μ_0	μ_1	c_1	μ_2	c_2	μ_3	c_3	μ_4	c_4
1	1	10	0.0	10.00	2.0	0.71	10.0	0.32	10.0	0.32
1	1	100	0.1	3.16	20.0	0.22	100.0	0.10	100.0	0.10
1	1	1000	1.0	1.00	200.0	0.07	1000.0	0.03	1000.0	0.03
1	2	10	0.0	10.00	2.0	0.71	10.0	0.32	10.0	1.05
1	2	100	0.1	3.16	20.0	0.22	100.0	0.10	100.0	1.00
1	2	1000	1.0	1.00	200.0	0.07	1000.0	0.03	1000.0	1.00
2	1	10	0.0	10.00	2.0	0.71	10.0	0.32	10.0	0.32
2	1	100	0.1	3.16	20.0	0.23	100.0	0.10	100.0	0.10
2	1	1000	1.0	1.00	200.0	0.10	1000.0	0.04	1000.0	0.04
2	2	10	0.0	10.00	2.0	0.71	10.0	0.32	10.0	1.05
2	2	100	0.1	3.16	20.0	0.23	100.0	0.10	100.0	1.01
2	2	1000	1.0	1.00	200.0	0.10	1000.0	0.04	1000.0	1.00

SUMMARY

This paper shows how the demand flows from the customers to the dealers to the DC's and to the supplier. The demand measures are the average monthly demand, the standard deviation and the coefficient of variation. The monthly demands at an average dealer are assumed as Poisson. The tables shows how the shape of the monthly demands can range from normal to exponential and beyond. Recall, when cov $\leq$ 0.33, the monthly demands could be shaped like a normal distribution, and when cov $\geq$ 1.00, the monthly demands are the lumpy type.

MODELING SUPPLY CHAIN NETWORK DESIGN AND PRODUCT RECOVERY PLANNING UNDER UNCERTAINTY

Nasr-Eddine Dahel, Graduate School of Int'l Policy and Management, Monterey Institute of International Studies, 460 Pierce Street, Monterey, CA 93940, 831-647-4602, edahel@miis.edu

ABSTRACT

We consider the design of a multi-echelon, closed-loop supply chain network and the planning of product recovery for remanufacturing under uncertainty. The network comprises a number of plants and disassembly centers of unknown locations (to be selected from a set of potential sites) and a number of customer zones and disposal sites at fixed locations. Demand for remanufactured products, return of used products, and the rate of product recovery are all considered as uncertain parameters. The system is modeled as a multi-period cost-minimizing stochastic programming problem to determine the location of plants and disassembly centers, and to specify, for each period in the planning horizon, the optimal remanufacturing/production quantities, the number of used products to collect and disassemble, the inventory of recovered cores to maintain as well as product flows between the various facilities in the supply chain.

Keywords: Closed-Loop Supply Chains, Reverse Logistics, Stochastic Programming.

INTRODUCTION

Closed-loop networks link together two distinct markets, namely a "disposer market" from which used products are collected, and a "reuse market" in which demand for remanufactured product exists (Fleischmann et al 2001). The intersection of these two heterogeneous markets produces mismatch between supply and demand. Availability of used products for recovery is less predictable than supply of new input materials in a traditional supply chain. Therefore, mismatch between supply and demand with respect to quantity and timing is more prevalent in closed-loop than in traditional supply chains.

Another major characteristic of closed-loop product recovery networks is the level of uncertainty about the quality of used products. In general, used product quality is not known beforehand and can, depending on the condition of the individual product, be subject to considerable variability. As a result, disassembly inspection and testing activities play an important role in transitioning the product from the disposer to the reuse market. The quantity of used products that may be reused, and the quantity to be disposed of, and hence the magnitude and destination of the various reverse flows can only be determined after disassembly and testing. Therefore the product recovery ratio is subject to uncertainty.

Remanufacturing can be carried out by a local manufacturer or an original equipment manufacturer (Pranab and Harry 2001). Considered in this paper is the latter case wherein the manufacturer remanufactures products from returned cores and other major components in parallel with the manufacturing of new products in the same facilities. In this environment, recovery networks are not commonly established from scratch but are designed using the existing set of plants and other logistics facilities. To this end, it is important to know which plants and disassembly centers to open and operate, and the number of units to process, store, and distribute out of them. Also, since capacity and recovery cost are facility-dependent, there is interest in determining whether it is economical to collect all returns and, by virtue of consequence, service all customer zone demands; and if not determining the appropriate level of collection remanufacturing and distribution of the recovery operation. Hence, facility and transportation decisions have to be integrated with recovery planning decisions so that material requirements, inventory levels, demand, and capacity constraints over the various stages of collection, disassembly, recovery, and disposition can be coordinated in a most economical way.

In the present work a multi-echelon logistics network model is developed for designing a closed-loop supply chain and planning product recovery for remanufacturing under conditions of uncertainty taken over a multi-period planning horizon. The supply chain structure consists of a number of plants and disassembly centers (to be selected from a set of potential locations) and a number of existing customer zones and disposal sites at fixed locations. Demand for remanufactured products, return of used products, and the rate of product recovery are all considered as uncertain parameters. Decisions provided by the cost-minimizing stochastic programming model determine the following: (1) the plants and disassembly centers to operate during the planning horizon, (2) the quantity to be produced at each plant and shipped to each customer zone in every time period, (3) the quantity of used products to be collected from each customer zone and shipped to each disassembly center in every period, (4) the quantity of reusable units each disassembly center ships to each plant in every time period, (5) the inventory of reusable units held at each disassembly center in every period, (6) the quantity of non-reusable units produced in each disassembly center and shipped to each disposal site in every time period.

MODEL DEVELOPMENT

The proposed model follows the closed-loop network structure shown in Fig. 1. We consider four types of facilities, namely plants where remanufacturing of the reusable units takes place, disassembly centers where the inspection and disassembly function of the used units is carried out, disposal sites where non-reusable units are disposed of, and customer zones in which remanufactured units are sold and from which used units are collected. Moreover, two outcomes are possible for the collected used units: recovery and disposal. Only a given fraction of the used units processed in the disassembly centers is deemed recoverable and therefore reusable during remanufacturing, the remaining units are considered non-reusable and thus disposable. We also consider two types of flows: forward and reverse flows. Forward flows represent shipments of remanufactured units from plants to customer zones. The reverse flows represent: (1) transportation of used units from customer zones to disassembly centers, (2) shipments of reusable units from disassembly centers to plants, and (3) transportation of non-reusable units from disassembly centers to disposal sites.

The following assumptions are postulated: (1) the supply chain facilities (plants, customer zones, disassembly centers, and disposal sites), already exist. (2) Demand for remanufactured products and supply of used products at customer zones are subject to uncertainties described by a given set of scenarios. (3) Plant production capacities, customer zone collection capacities, and disassembly center capacities are known. (4) A product recovery ratio determines the number of reusable units resulting from disassembly and inspection of used units. This ratio is common to

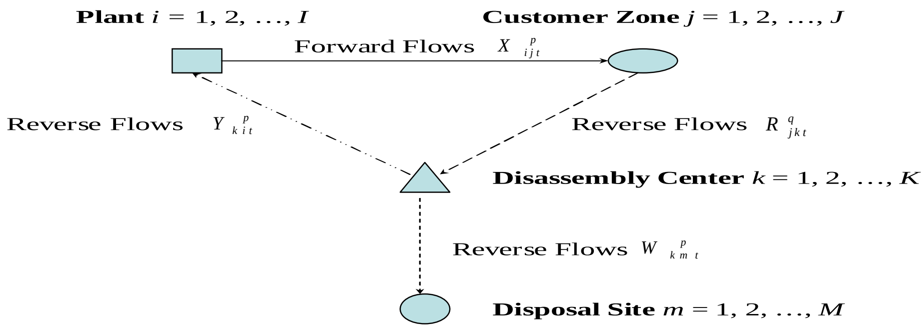


Fig. 1 Closed-Loop Network

all units regardless of the disassembly center they are processed in and the customer zone they are collected from. (5) Inventory of reusable units is held at the disassembly centers. (6) A minimum proportion of the reusable units recovered within a given disassembly center in a time period must be shipped out of that center in that period. This assumption is designed in such a way to ensure an adequate inventory turnover at each center and thereby reduce obsolescence of the reusable unit inventory in the supply chain. (7) Disposal sites have unlimited capacities.

Notation

Primary Sets and indices

$I =$ Set of plants in which the product may be remanufactured, $i \in I$;

$K =$ Set of disassembly centers in which returns may be processed, $k \in K$;

$J =$ Set of customer or demand points, $j \in J$;

$M =$ Set of disposal sites, $m \in M$;

$T =$ Set of time periods, $t \in T$;

$P =$ Set of demand of scenarios, $p \in P$;

$Q =$ Set of product return scenarios, $q \in Q$;

$R =$ Set of product recovery scenarios, $r \in R$;

Supply/Demand Parameters

Demand uncertainty is represented by a set of p scenarios, each having a probability of occurrence equal to π^p. Likewise, used product return uncertainty is represented by a set of q scenarios, along with their corresponding probability π^q. Uncertainty of used product quality is, in a similar manner, represented by a set of r scenarios, each having an occurrence probability of π^r. Also, let π^{pq} define the joint probability of the concurrent occurrence of demand scenario p and return scenario q. These probabilities will generally satisfy the conditions that $\sum_{p\in P}\pi^p = \sum_{q\in Q}\pi^q = \sum_{r\in R}\pi^r = \sum_{p\in P}\sum_{q\in Q}\pi^{pq} = 1$. Furthermore let

$D^p_{jt} =$ Customer zone j demand under scenario p during time period t;

$d_{jt}^{q} =$ Customer zone *j* used product return under scenario q during period *t*;

$S_i =$ Plant *i* production capacity per period;

$S_j =$ Customer zone *j* used product collection/storage capacity per period;

$S_k =$ Center *k* disassembly capacity per period;

$m_k =$ Minimum shipment requirement out of an open center *k* to plants per period;

$\theta_k =$ Center *k* inventory storage capacity;

$\lambda^r =$ Product recovery ratio during disassembly under scenario *r*;

Cost Parameters

$f_i =$ Fixed cost for opening and operating plant *i*;

$f_k =$ Fixed cost for opening and operating center *k*;

$c_{ki} =$ Per unit remanufacturing cost at plant *i* using materials sourced from processing center *k*; this cost comprises unit production cost at plant *i*, plus unit transportation cost from center *k* to plant *i*.

$c_{jk} =$ Per unit disassembly cost at center *k* of a used product collected at demand point *j*; this cost includes unit collection cost at point *j*, transportation cost per unit from *j* to *k*, and disassembly cost of a unit at *k*.

$c_{km} =$ Per unit disposal cost at site *m* of a non-recoverable unit processed at center *k*. This cost includes disposal cost of a unit at *m* plus unit transportation cost from *k* to *m*;

$h_k =$ Per unit per period inventory holding cost of a disassembled unit in inventory at center *k*;

$p_j =$ Unit penalty cost for not collecting returns of customer zone *j*;

$g_j =$ Unit penalty cost of not serving demand of customer zone *j*. g_j could be quantified by taking the relative importance of different markets into account; alternatively it could be related to the cost of meeting demand by resorting to external suppliers.

$t_{ij} =$ Unit transportation cost of a product from plant *i* to customer zone *j*;

Decision Variables

$X_{ijt}^{p} =$ Forward flow: units shipped from plant *i* to customer zone *j* under demand scenario *p* in period *t*;

$R_{jkt}^{q} =$ Reverse flow: units of used product shipped from customer zone *j* and to center *k* under return scenario q in period *t*;

$Y_{kit}^{p} =$ Units of cores shipped from center *k* to plant *i* under demand scenario p in period *t*. This flow reflects, by the same token, the number of units produced in plant *i* out of cores sourced from center *k* to respond to demand scenario p in period *t*;

$W_{kmt}^{q} =$ Units shipped from disassembly center *k* to disposal site *m* under return scenario q in period *t*;

$B_{jt}^{p} =$ Units of unsatisfied demand at customer zone *j* under scenario *p* in period *t*;

I_{kt}^{pq} = Disassembled units held in inventory at center k at the end of period t under demand scenario p and return scenario q;

U_{jt}^{q} = Uncollected used units at customer zone j under scenario q at the end of period t;

$Z_i = \begin{cases} 1, \text{ if product is produced in plant } i; \\ 0, \text{ otherwise.} \end{cases}$ $V_k = \begin{cases} 1, \text{ if used product is disassembled in center } k; \\ 0, \text{ otherwise.} \end{cases}$

The complete mixed integer programming model can now be formulated as follows.

$$\text{Min} \sum_{i\in I} f_i Z_i + \sum_{k\in K} f_k V_k + \sum_{t\in T}[\sum_{p\in P} \pi^p (\sum_{i\in I}\sum_{k\in K} c_{ki} Y_{kit}^p + \sum_{j\in J} g_j B_{jt}^p + \sum_{i\in I}\sum_{j\in J} t_{ij} X_{ijt}^p) + \sum_{q\in Q}\pi^q (\sum_{j\in J} p_j U_{jt}^q + \sum_{j\in J}\sum_{k\in K} c_{jk} R_{jkt}^q + \sum_{k\in K}\sum_{m\in M} c_{km} W_{kmt}^q) + \sum_{p\in P}\sum_{q\in Q}\pi^{pq}(\sum_{k\in K} h_k I_{kt}^{pq})] \quad (1)$$

Subject to:

$$\sum_{j\in J} X_{ijt}^p \le S_i Z_i, \qquad i\in I,\ p\in P,\ t\in T; \quad (2)$$

$$\sum_{i\in I} X_{ijt}^p + B_{jt}^p = D_{jt}^p, \qquad j\in J,\ p\in P, t\in T; \quad (3)$$

$$\sum_{k\in K} Y_{kit}^p = \sum_{j\in J} X_{ijt}^p, \qquad i\in I,\ p\in P, t\in T; \quad (4)$$

$$I_{k,t-1}^{pq} + \sum_{r\in R}\sum_{j\in J}\pi^r \lambda^r R_{jkt}^q = \sum_{i\in I} Y_{kit}^p + I_{kt}^{pq}, \qquad k\in K,\ p\in P, q\in Q, t\in T; \quad (5)$$

$$I_{kt}^{pq} \le \theta_k V_k, \qquad k\in K,\ p\in P, q\in Q, t\in T; \quad (6)$$

$$\sum_{k\in K} R_{jkt}^q + U_{jt}^q = d_{jt}^q, \qquad j\in J,\ q\in Q, t\in T; \quad (7)$$

$$\sum_{k\in K} R_{jkt}^q \le S_j, \qquad j\in J,\ q\in Q,\ t\in T; \quad (8)$$

$$\sum_{j\in J} R_{jkt}^q \le S_k V_k, \qquad k\in K,\ q\in Q,\ t\in T; \quad (9)$$

$$\sum_{i\in I} Y_{kit}^p \ge m_k V_k, \qquad k\in K,\ p\in P,\ t\in T; \quad (10)$$

$$\sum_{m\in M} W_{kmt}^q = \sum_{r\in R}\sum_{j\in J}\pi^r(1-\lambda^r) R_{jkt}^q, \qquad k\in K,\ q\in Q,\ t\in T; \quad (11)$$

$$V_k, Z_i = \{0, 1\}, \qquad i\in I,\ k\in K; \quad (12)$$

$$X_{ijt}^p, R_{jkt}^q, Y_{kit}^p, W_{kmt}^q, I_{kt}^{pq}, B_{jt}^p, U_{jt}^q \ge 0, \qquad i\in I, j\in J, k\in K, m\in M,\ p\in P, q\in Q, t\in T; \quad (13)$$

Constraints (2) specify that the total flow out of plant i, and thereby the total number of units produced at plant i, during period t must be less than or equal to that plant production capacity if the product is produced in such a plant; and must be equal to zero otherwise. Constraints (3) ensure product flow balance between forward product flows into demand point j, and demand requirement for demand point j at time period t, and account for the possibility of unsatisfied demand at that demand point. Unsatisfied demand occurs when not enough returned cores are collected or when product demand is greater than production and/or disassembly capacities. Equation (4) is a material balance constraint ensuring that the sum of the quantities going into a plant i (or reverse flow) equals the sum of the quantities coming out of that plant (or forward flow) in every time period. Constraints (5) ensure product flow balance between inventory, recovery, and shipment of disassembled units at disassembly center k in time period t. Inventory, determined in this case on the basis of the expected number of recovered parts taken over all possible recovery scenarios, may be carried to provide better customer service or to satisfy forecasted demand that exceed production capacities in future time periods. Constraints (6) specify that the total number of disassembled units stored in inventory at center k in period t cannot be larger than the inventory storage capacity of that center. Constraints (7) ensure product flow balance between collection of returns, and forecasted returns at demand point j in time period t, and by the same token accounts for any uncollected returns at that demand point and time period. Constraints (8) require that the total number of units collected at demand point j in time period t to be less than the specified collection/storage capacity of that demand point. Equation (9) requires the total flow into center k, and thereby the total number of units processed at such a center, during period t to be less than or equal to that center processing capacity if returns are disassembled in such a center; and must be equal to zero otherwise. Used units collected from customer zones are assumed to be processed within the same time period in which they are collected. Constraints (10) require that the total flow out of any open center k in period t meets the minimum output requirement for that center. Constraints (11) specify the number of non-recoverable units transported from disassembly center k to disposal sites. Constraints (12) and (13) respectively ensure integrality and nonnegativity on the decision variables.

The objective function (1) minimizes the *expected value* of the total multi-period cost of production, collection, disassembly, disposal, inventory, and transportation of the network taken over all the given scenarios. The components of the objective function may be described as follows:

Fixed costs at plants and disassembly centers over the entire planning horizon $= \sum_{i \in I} f_i Z_i + \sum_{k \in K} f_k V_k$.

Multi-period variable production cost at the plants $= \sum_{p \in P} \pi^p (\sum_{t \in T} \sum_{i \in I} \sum_{k \in K} c_{ki} Y_{kit}^p)$.

Inventory costs at processing centers $= \sum_{p \in P} \sum_{q \in Q} \pi^{pq} (\sum_{t \in T} \sum_{k \in K} h_k I_{kt}^{pq})$.

Penalty cost of unsatisfied demand at customer zones $= \sum_{p \in P} \pi^p (\sum_{t \in T} \sum_{j \in J} g_j B_{jt}^p)$.

Penalty cost of uncollected returns at customer zones $= \sum_{q \in Q} \pi^q (\sum_{t \in T} \sum_{j \in J} p_j U_{jt}^q)$.

Collection, transportation, and processing costs of used units $= \sum_{q \in Q} \pi^q (\sum_{t \in T} \sum_{j \in J} \sum_{k \in K} c_{jk} R_{jkt}^q)$.

Disposal and transportation costs of non-recyclable units $= \sum_{q \in Q} \pi^q (\sum_{t \in T} \sum_{k \in K} \sum_{m \in M} c_{km} W_{kmt}^q)$.

Transportation cost from plants to customer zones $= \sum_{p \in P} \pi^p (\sum_{t \in T} \sum_{i \in I} \sum_{j \in J} t_{ij} X_{ijt}^p)$.

References available upon request from the author.

A SUPPLY CHAIN STRATEGIC PLANNING FRAMEWORK

Matthew J. Liberatore, Villanova University, 610-519-4390, matthew.liberatore@villanova.edu
Tan Miller, Rider University, 973-590-4638, tanjean@verizon.net

ABSTRACT

Developing and implementing a supply chain strategy is essential for maintaining competitive position. As part of the strategy development process, firms must select projects that best achieve strategy. In this paper we present a framework that a firm can employ to select those projects which will best support the firm's supply chain objectives and strategies. A multi-criteria decision-making technique, such as the Analytic Hierarchy Process, can be used to support this strategic planning process. An example is presented to illustrate the proposed approach.

INTRODUCTION

We present a framework that a firm can employ to select those projects which best support the firm's supply chain objectives and strategies. This framework offers a well-defined process by which firms can select those projects that will yield the greatest contributions towards achieving their objectives and executing their strategies. Our framework employs the analytic hierarchy process (AHP), a multi-criteria decision-making technique, to guide a firm through its strategic planning process. This guidance includes a methodology for developing project priorities subject to budget constraints and other organizational requirements such as the need to allocate resources across different functional areas and strategies.

LITERATURE REVIEW

A supply chain strategy should be selected based on the characteristics of the product and markets served. Fisher [3] categorized products as either functional or innovative, depending on the predictability of demand, size of margins, length of product life cycles, and level of product variety. Functional products face a stable demand process and low margins, and so their supply chains should focus on cost efficiency. Innovative products face unpredictable demand but have higher margins, and so a responsive strategy is desired, and can be achieved using mass customization or product postponement.

Characteristics of the supply processes, including the number of supply sources and process changes, and variable or stable yields can also be related to the choice of a supply chain strategy. Lee [5] proposes that products having predictable demand and stable supply sources require an efficient supply chain strategy, while those having unpredictable demand and stable supply sources require a responsive strategy [3]. However, if supply processes are evolving Lee [5] proposes two additional strategies, depending on the level of demand uncertainty. When demand is predictable and the supply process is evolving, Lee proposes a risk hedging strategy that pools resources to share risks in supply disruption. An agile supply chain is needed when demand is unpredictable and supply is evolving. It must be flexible enough to respond to changing levels of demand, while trying to minimize the supply disruptions. Autry, Zacharia, and Lamb [1] developed a classification scheme for logistics strategy types based on the results of a survey of managers and executives knowledgeable about their firm's logistics strategies. Using cluster analysis, two strategies were determined: functional logistics and externally-oriented logistics. Those firms taking a functional logistics strategy focus on inventory and order management, order processing, procurement, and storage. The underlying measurement items emphasize efficiency and cost management, analysis and control. Those firms employing an externally-oriented strategy are most concerned with ensuring that their logistics activities are compatible with the needs of their trading partners and/or directed toward stewardship of the overall business environment (e.g., green or socially responsible logistics). Firms adopting this strategy focus more on coordination and collaboration, social responsibility, strategic distribution planning and technology and information systems

Defee and Stank [2] discuss how the strategy-structure-performance paradigm [4], [7] can be extended from a single-firm to a supply chain environment that includes the firm's channel partners. The presumption is that competition has moved from the firm to the supply chain level. Therefore, inter-firm coordination is required to achieve a supply chain strategy. Sadler and Hines [8] discuss an integrated supply chain planning process that attempts to involve all partner companies or intermediaries into the process. They describe the application of their strategic operations and logistics process (SOLP) to the Australian meat processing industry [8] and at a heavy fabrication business in Australia [9]. However, the approach followed by many manufacturers is to develop a supply chain strategic plan internally, while considering the relationships with their partners during the planning process.

A review of the supply chain strategy and planning literature has informed our research on the issues that must be addressed in the supply chain planning process.

PLANNING FRAMEWORK

We use an example firm, the ABC Manufacturing Company, to illustrate how supply chain objectives, strategies and project selection process can be integrated into one aligned, comprehensive approach. ABC produces a wide variety of consumer goods and distributes its products to a broad array of retailers who then sell these goods to consumers. Recently ABC has also implemented its own direct-to-consumer channel of distribution, an effort still in initial stages of development. ABC's supply chain organization is just now entering its annual strategic planning process.

To develop and organize its supply chain objectives and strategies, ABC first employs the mission, objective and strategy (MOS) approach to strategic planning [6]. The MOS approach represents a planning framework in which a firm defines its overall mission (*step 1*), establishes objectives to support its mission (*step 2*), and then develops strategies to support its objectives (*step 3*). Similarly, a major organization within a firm, such as Supply Chain, can also adapt the MOS approach for its functional area. ABC's supply chain organization utilizes the MOS methodology to form the first three tiers of its supply chain planning framework. The first level – ABC's overall supply chain mission – is stated as:

> <u>**MISSION**</u>**:** ***Position ABC's supply chain to support the overall firm objectives as efficiently and effectively as possible***

In the second stage ABC evaluates its supply chain from both internal and external perspectives, and based on this evaluation, defines a set of objectives designed to achieve its mission. The ABC team reviewed a wide range of supply chain functional capabilities, and based upon this review, defined explicit performance objectives in the following areas:

> **SUPPLY CHAIN OBJECTIVES**
>
> ***i.*** ***Regulatory Compliance***
> ***ii.*** ***Cost Efficiency***
> ***iii.*** ***Delivery Effectiveness***
> ***iv.*** ***Flexibility***
> ***v.*** ***Security***

We note that ABC (and any firm utilizing this process) should formulate precise objectives against which it can accurately evaluate itself in these defined functional areas. However, for illustrative purposes, we limit our discussion of ABC's objectives to this simple description of the general areas of objectives.

In the third planning stage, ABC develops strategies designed to meet those objectives posed in the second stage. Here ABC's managers consider and address the needs of all of the major stakeholders and constituents of the firm and its supply chain. ABC's supply chain management addressed six key areas where it believed that well-defined and "measurable" strategies were necessary to achieve these five objectives. ABC determined that it required strategies to address the following major stakeholders and constituents of its supply chain:

> **SUPPLY CHAIN STRATEGIES**
>
> ***i.*** ***The Consumer (CONS)***
> ***ii.*** ***Its Customers (CUST)***
> ***iii.*** ***Its Governmental and Societal Responsibilities (G&S)***
> ***iv.*** ***Its Financial Well-being (FIN)***
> ***v.*** ***Its Suppliers and Partners (S&P)***
> ***vi.*** ***Its Internal Operations and Employees (OPS&E)***

Again, for the illustrative purposes of this paper, we will not describe the details of the explicit strategies that ABC would have developed. Instead, we simply note the general areas of strategy that address all of ABC's stakeholders and constituents.

The mission, objectives and strategies defined by ABC's supply chain management provide a framework for the organization to move forward. With its strategies defined, now ABC must select and initiate individual projects that will facilitate the successful implementation of its strategies. For ABC, this next process consists of first identifying a set of potential supply chain strategic projects, and then selecting the actual projects it will pursue among the candidate projects proposed. The list of potential supply chain projects and the strategy they support are shown in Figure 1. Figure 2 displays the complete framework.

No.	Project Description	Strategy Supported
1	Streamline Existing Direct To Consumer Delivery and Internet Ordering Services	CONS
2	Integrate POS (point of sale) Data As Input To The Current Forecasting and Production Scheduling Processes	CUST
3	Establish and Enhance Customer Relationship Management Services (CRM)	CUST
4	Implement A Customer Segmentation Program Based On Customer Profitability & Activity Based Costing (ABC) Analyses	CUST
5	Expand Supplier Relationship Management (SRM) Programs	S&P
6	Evaluate and Benchmark Current Security Processes and Procedures on Supply Chain, and Recommend Enhancements	G&S
7	Integrate Current Individual Country Sales & Operations Planning Processes Into One Global S&OP Process	OPS&E
8	Develop And Implement Plan to Optimize Use of RFID In Supply Chain and use of RFID vs. use of Bar Codes	OPS&E
9	Evaluate Current Use Of Third Party Logistics (3PL) Firms And Recommend If Firm Should Increase Use Of 3PLs	FIN
10	Develop and Implement Labor Productivity Measurement Systems For plant and warehouse operations	OPS&E
11	Expand Current Import / Export Compliance IT Capabilities	G&S
12	Benchmark Firm's Current Carbon Footprint and Develop Long-Term Green Manufacturing Operations Plan	G&S

Figure 1. Project Descriptions and Strategies Supported

ANALYSIS

To determine the set of projects that best supports the supply chain mission, a multicriterion method such as the Analytic Hierarchy Process could be used [10]. An AHP analysis uses judgments in the form of pairwise comparisons to measure the impact of items on one level to the next higher level. At each level, the pairwise comparisons are organized into a matrix and the weights of the items being compared are determined by computing the maximum eigenvalue of the matrix. A weighted averaging approach is used to combine the results across levels of the hierarchy to compute a final weight for each alternative.

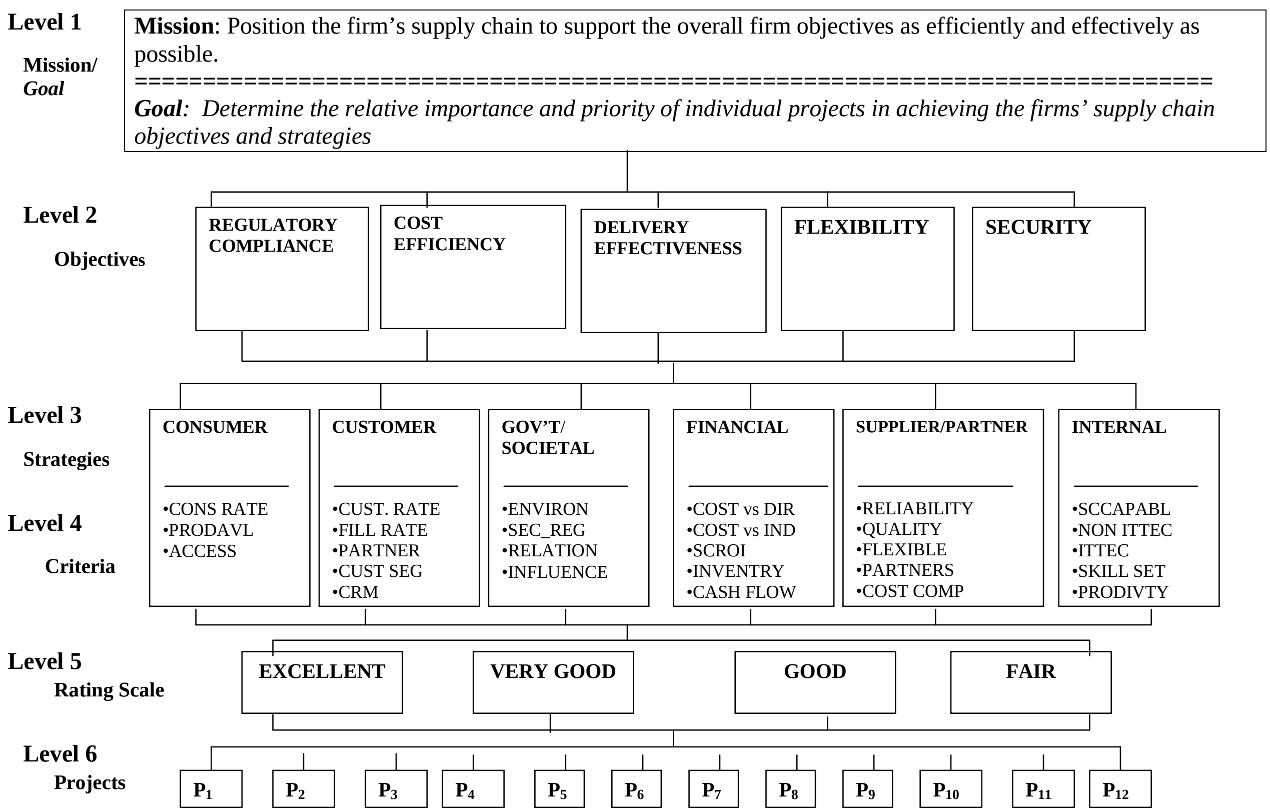


Figure 2. Supply Chain Strategic Planning Framework

As applied to Figure 2, the planning team would determine the relative importance of the objectives in achieving the mission. For example, how much more important is regulatory compliance than cost efficiency (or vice versa) in achieving the mission? Next, the team would determine the relative importance of the strategies in achieving each objective. For example, with respect to the cost efficiency objective, how much more important is the consumer strategy as compared to the customer strategy? For each strategy, the criteria must be compared so that the weights needed to evaluate the projects can be established. For each criterion a rating scale is then established, such as excellent, very good, good, and fair. Weights are developed for each of these ratings, and so projects can be scored. For example, since project three supports the Customer strategy, it must be rated with respect to the five Customer criteria (CUSTRATE, FILLRATE, PARTNER, CUSTSEG, CRM). Each of the criteria has specific definitions. For example, the ratings scale for FILLRATE is: *Excellent* = project will move firm's line item fill rate to customers to 99% or higher from its current 95% level, *Very Good* = 98% to 99% fill rate, *Good* = 96% to 98% fill rate, and *Fair* = 95% to 96% fill rate. Those projects having the highest scores would be funded, subject to budget availability.

CONCLUSIONS

In this paper, we have developed a comprehensive framework to support a firm's supply chain strategic planning activities and its project selection process. Our framework combines two well established methodologies, the MOS (for strategic planning) and the AHP (for multi-criteria decision-making/project selection) to formulate one aligned planning process. In particular, this framework and methodology allows a firm to ensure that the individual supply chain projects that it pursues support and align with its strategies and objectives. The framework provides a firm with an approach to first establish its mission, objectives and strategies. Then this framework allows a firm's management to qualitatively and quantitatively select and prioritize those projects that will best fulfill its strategies and objectives. In the corporate supply chain planning process, where difficult and often budget-constrained decisions must be made on which projects to implement, our framework offers the guidance necessary to assure aligned decision-making.

REFERENCES

[1] Autry, C. W., Zacharia, Z. G., Lamb, C. W., " A Logistics Strategy Taxonomy," *Journal of Business Logistics*, 29(2), 2008, 27 – 51.

[2] Defee, C. C., Stank, T. P., "Applying the Strategy-Structure-Performance paradigm to the Supply Chain Environment," *The International Journal of Logistics Management*, 16(1), 2005, 28 - 50.

[3] Fisher, M. L., "What Is the Right Supply Chain for Your Product," *Harvard Business Review*, 75(2), 1997, 105 – 116.

[4] Galbraith, J. R., Kazanjian, R. K., *Strategy Implementation: Structure, Systems, and Process,* West Publishing Company, St. Paul, MN, 1986.

[5] Lee, H. L., "Aligning Supply Chain Strategies with Product Uncertainties," *California Management Review*, 44(3), 2002,105 – 119.

[6] Liberatore, M. J., Monahan, T.F., Stout, D.E., "A Framework For Integrating Capital Budgeting Analysis With Strategy," *Engineering Economist*, 38(1), 1992, 31-43.

[7] Miles, R. E., Snow, C. C., *Organizational Strategy, Structure, and Process*, McGraw-Hill, New York, NY, 1978.

[8] Sadler, I., Hines, P., "Strategic Operations Planning Process for Manufacturers with a Supply Chain Focus: Concepts and a Meat Processing Application," *Supply Chain Management: An International Journal*, 7(4), 2002, 225 – 241.

[9] Sadler, I., Gough, R., "Applying a Strategic Planning Process to Several Supply Chain Planning Partners," *Journal of Manufacturing Technology Management*, 16(8), 2005, 890 – 908.

[10] Saaty, T. L., *The Analytic Hierarchy Process*, Pittsburgh, PA: RWS Publications, 1996.